ABSTRACT

Yang, Tung-Sheng. Performance Analysis of Adaptive Transmission Aided by Long Range Channel Prediction for Realistic Single- and Multi-Carrier Mobile Radio Channels. (Under the direction of Dr. Alexandra Duel-Hallen)

Reliable adaptive transmission for frequency selective mobile radio systems is addressed. In particular, we investigate adaptive channel loading for wireless Orthogonal Frequency Division Multiplexing (OFDM) systems and adaptive modulation aided by observations of another carrier (e.g, Frequency Division Duplex (FDD) channels). Adaptive transmission techniques, where the modulation size, coding rate, or other signal transmission parameters are dynamically adapted to the changing channel conditions, have recently emerged as powerful tools for increasing the data rate and spectral efficiency for wireless system. However, reliable adaptive transmission requires long-range prediction (LRP) of future channel state information (CSI) due to the variation of the wireless channel, which results in different channel conditions between the time of data transmission and the time of the channel estimation. We derive the minimum mean-square-error (MMSE) long range channel prediction method that utilizes the time and frequency domain correlation function of the Rayleigh fading channel. Since the channel statistics are usually unknown, reduced complexity robust prediction methods that can converge rapidly to the theoretical MMSE and do not require the knowledge of correlation functions are developed for OFDM channels and systems aided by observations of another carrier. Statistical model of the prediction error that depends on the frequency and time correlation is developed and is used in the design of reliable adaptive modulation methods.

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
2004		**00-00-2004 to 00-00-2004**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Performance Analysis of Adaptive Transmission Aided by Long Range Channel Prediction for Realistic Single- and Multi-Carrier Mobile Radio Channels	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **North Carolina State University,Department of Electrical Engineering,Raleigh,NC,27695**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
The original document contains color images.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**		**133**	

A standard sum-of-sinusoids Rayleigh fading channel model and a novel physical model based on the method of images augmented with diffraction are employed to test the prediction algorithm. This physical model can generate non-stationary datasets to test both the LRP and its application in adaptive transmission schemes. It is demonstrated that this physical model generates realistic datasets that closely resemble measured data, and the results of the LRP for the physical model and measured data are similar, and differ significantly from those produced for the Jakes model. We use this model to produce different scenarios to classify typical and challenging cases to test the performance of the proposed prediction algorithm. These cases are more difficult to identify with the measured data. Moreover, we examine the dependency of the correlation between two different carrier frequencies on the variation of the root mean square (*rms*) delay spread and investigate the limits on the adaptation rate in adaptive transmission systems aided by observations of another carrier. Thus, the physical model allows to test robustness and to determine practical constraints of the proposed adaptive transmission methods.

Performance Analysis of Adaptive Transmission Aided by Long Range Channel Prediction for Realistic Single- and Multi-Carrier Mobile Radio Channels

by
Tung-Sheng Yang

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

ELECTRICAL ENGINEERING

Raleigh
2004

APPROVED BY:

Hamid Krim	Brian L. Hughes
Hans Hallen	Alexandra Duel-Hallen *Chair of Advisory Committee*

To my family, friends,
and my beloved An-Chun

BIOGRAPHY

Tung-Sheng Yang was born in Taichung, Taiwan on March 3rd, 1976. He received the Bachelor of Science degree in Electrical and Electronics Engineering from Tsing Hua University, Hsinchu, Taiwan in August 1998. He started his graduate studies at North Carolina State University, Raleigh NC in Electrical and Computer Engineering Department in Fall 1998. In 2000, he received the Master of Science degree. He was enrolled in the Ph.D. program and was a teaching assistant in winter 2000 in the same department. He became a research assistant in fall 2000. Since then, he has been a research assistant at Center for Advanced Computing and Communications (CACC).

The author intends to pursue a career in the field of digital and wireless communications.

ACKNOWLEDGMENTS

I am indeed grateful to my advisor, Dr. Alexandra Duel-Hallen for her supervision, support and encouragement through my academic research. I truly appreciate her enthusiastic help in suggesting research topics, revising my technical papers, and preparing my presentations.

I would like to thank Dr. Hans Hallen for his time as my committee member, and for his help, instructions, and suggestions for applying his physical model of fading channels on my research work. I would also like to thank Dr. Brian L. Hughes and Dr. Hamid Krim for their time to serve as my advisory committee.

I express my appreciation to my good friends and colleagues, with whom I had many fruitful conversations about my research work: Dr. Shengquan Hu, Dr. Ayman El-Ezabi, Dr. Zhihong Hong, and Ming Lei.

I thank Center for Advanced Computing and Communications (CACC), the Department of Electrical and Computer Engineering at North Carolina State University, the National Science Foundation, and the Army Research Office for their support and help of this work.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUTION

1.1 Background and Motivation

Wireless communications is experiencing an explosive growth rate due to enabling technologies which permit wide spread deployment. Continued increase in demand for all types of wireless services such as voice, data, and multimedia is fueling the need for high capacity and data rate. This high demand for wireless communications services requires increased system capacities. Therefore, new transmission methods and service have been enthusiastically adopted by people throughout the world. Particularly during the past ten years, the mobile radio communications industry has grown by orders of magnitude, fueled by digital and RF circuit fabrication improvement, new large-scale circuit integration, and other miniaturization technologies which make portable radio equipment smaller, cheaper, and more reliable. Digital switching techniques have facilitated the large scaled deployment of affordable, easy-to-use radio communication networks. These trends will continue at an even greater pace during the next decade.

It is well known that the fundamental limitation of wireless systems is constituted by their time-variant channel fading, which results in dramatic signal-to-noise (SNR) fluctuations observed in both time and frequency. The traditional wireless cellular systems are designed to provide good transmission quality for the worst channel conditions. As a result, SNR that are much larger than the target are achieved over a large portion of the cellular coverage area and transmission time, which leads to inefficient utilization of the full channel capacity. In addition, the integration of the voice and data transmission has caused

different quality of service (QoS) requirements over wireless systems. Voice transmission requires a certain minimum SNR and can only tolerate a latency of up to 100ms. On the other hand, data is more flexible. Data flow may be increased in good channels to boost the throughput and conversely, reduced in poor conditions in exchange for a lower bit error rate (BER).

Motivated by the above mentioned performance limitations of fixed-mode transceiver, adaptive transmission (AT) techniques, where the modulation level, coding rate, and other signal transmission parameters are dynamically adapted to the changing channel conditions without sacrificing the BER, have recently emerged as powerful tools for increasing the data rate and spectral efficiency and have attracted considerable research interests in the past decades [21,22,57,64]. The AT techniques can be applied to single-carrier [16,21,22,24,31,57], multi-carrier Orthogonal Frequency Division Multiplex (OFDM) [11,10,26,36] as well as Code Division Multiple Access (CDMA) [1,41] transceivers. Besides, it can be incorporated in transmit antenna diversity systems [30,32]. The goal is to communicate over hostile mobile channels at a higher integrity or higher throughput, than conventional fixed-mode transceivers. A number of existing wireless systems already support some grade of adaptivity and future research is likely to promote these principles further by embedding them into the already existing standards. The range of various existing solutions that have found favor in already operational standard systems was summarized in [40] by Nanda et al.

The AT techniques generally require accurate channel state information (CSI) that can be acquired from different sources. If the communication between the two stations is bi-direction and the channel can be considered reciprocal, as, for example, in time division

duplex (TDD) systems, then each station can estimate the channel quality on the basis of the received symbols and adapt the parameters to this estimation. This is called open-loop adaptation [57]. If the channel is not reciprocal, the receiver has to estimate channel quality from feedback resulting in closed-loop adaptation. The feedback load consumes power and bandwidth, and the fed back CSI needs to be quantized resulting in degraded performance. Note that for many adaptive transmission applications (e.g., selective transmitter diversity or adaptive modulation [31]), it is not necessary to feed back the actual fading coefficient. It is sufficient to send to the transmitter just the antenna selection or modulation index bits derived from the estimates of predicted values at the receiver. The feedback delay, overhead, channel estimation and CSI quantization errors, and processing delay degrade the performance of adaptive modulation, especially in rapidly time variant fading. Even in open-loop channels, current CSI is not sufficient since future channel conditions need to be known to adapt transmission parameters. To realize the potential of adaptive transmission methods, the channel variations have to be reliably predicted at least several milliseconds ahead.

Recently, many researchers have developed techniques to predict the near-term behavior of the mobile channel. In [20], the multivariate adaptive regression splines (MARS) model was used to capture the dynamics for predicting parameters of wideband fading channels several millisecond ahead for fast vehicle speed. The sub-space based [46] root-MUSIC [27] method and ESPRIT [3] type algorithm are employed to estimate the power spectrum that constitutes the fading process. Then these sinusoids are extrapolated to predict future samples. In addition, a novel long-range prediction (LRP) method was proposed in [16,18,19,31]. This algorithm employs an autoregressive (AR) model to characterize the fading channel and computes the minimum mean-square-error (MMSE) estimate of a future-

fading coefficient based on a number of past observations. The advantage of this algorithm relative to conventional methods is due to its low sampling rate (on the order of twice the maximum Doppler shift and much lower than the data rate), which results in longer memory span and further prediction into the future for a fixed filter length. The low sampling rate also results in reduced feedback rate for closed-loop AT systems. All these methods were examined by synthetic and measured data, and it was concluded that reliable prediction is feasible at least one wavelength into the future (or 10 ms ahead for vehicle the vehicle speed of 68 miles per hour and the carrier frequency of 1 GHz). Since fades occur at least half a wavelength apart, this prediction capability means that enough information about the future deep fade can be forecasted to adjusting transmission modes in an optimal sense.

In this thesis, we focus on the adaptive modulation (AM) and power control (PC) aided by the long-range prediction technique on single- and multi-carrier systems. The LRP method was modified to predict not only in time but also in frequency domain. Since the channel statistics are time-variant, adaptive LRP techniques using least mean square (LMS) and recursive least-squares (RLS) algorithm are also developed.

For the single carrier system, we concentrate on the scenario where we observe a received uplink signal at the carrier frequency f^1 and attempt to predict the downlink signal at the carrier frequency f^2 without feedback from the mobile. Alternatively, a signal at frequency f^1 can be fed back and a signal at adjacent frequency f^2 is predicted without feedback. To accomplish this prediction, the predicted samples must be sufficiently correlated with the observations in both time and frequency. This technique can be applied in correlated uplink and downlink channels as in Frequency Division Duplex (FDD) systems, in orthogonal frequency division multiplexing (OFDM) systems (where narrow correlated sub-

channels are employed) or other wideband systems to reduce feedback and overhead requirements.

For the multicarrier system, we developed LRP for Orthogonal Frequency Division Multiplexing (OFDM) [5,63] systems and investigated the performance of adaptive bit and power loading for the OFDM system. OFDM has been proposed for use in high-speed wireless data applications due to its relatively simple receiver structure compared to single carrier transmission in frequency selective fading channel. In OFDM, the frequency band is divided into narrow subcarriers, and data bits are multiplexed onto these subcarriers. The systems eliminate the need for an equalizer by greatly reducing the intersymbol interference (ISI), and have been extremely popular in Digital Subscribe Line (DSL), digital audio and television broadcasting (DAB [61] and DVB) in Europe and wireless Internet access.

The fading channel is characterized as superposition of several scattered components. The prediction accuracy of the LRP algorithm is determined by the rate of change of amplitude, frequency and phase of each path [16,18,33,34]. However, the standard Jakes channel model [35] or a stationary random process description does not capture the variation of these parameters. To validate the LRP algorithm, a novel physical channel modeling based on the method of images and augmented with diffraction is proposed in [16,33,34]. This physical model can generate non-stationary datasets to test both the LRP and its application in adaptive transmission scheme. It is demonstrated in [16,33,34] that this physical model generates datasets that closely resemble measured data, and results of the LRP for the physical model and measured data are similar, and differ significantly from those produced for the Jakes model. Thus, we have demonstrated that the proposed physical model is realistic. In addition, this model's insights allow classification of scenarios into typical and

challenging cases for testing the algorithm. These scenarios are more difficult to identify with the measured data. Thus, the physical model allows us to test robustness and to determine practical constraints of the proposed adaptive transmission methods.

1.2 Outline of the Thesis

This thesis focuses on the investigation of the single- and multicarrier-adaptive transmission systems, physical mobile channel modeling and the long-range channel prediction algorithm in the time- and frequency domain. The thesis outline is as follows:

In chapter 2, we first briefly review the fading channel phenomenon over narrow band signal and develop a sum-of-sinusoids multipath fading channel model. We derive the correlation functions over time and frequency domain based on this model. Subsequently, we employ the standard Jakes model, a modified Jakes model, and a novel physical model to generate fading channels for numerical system simulations. We also discuss the limitation of the Jakes model and the properties of the physical model.

In chapter 3, we investigate the applications of the adaptive modulation and adaptive power control method. We discuss principles and possible scenarios where adaptive modulation can be applied, and then we review the fading channel capacity, which serves as the theoretical performance upper bound. Next, we present our adaptive modulation method utilizing multilevel quadrature amplitude modulation (MQAM) in conjunction with power control methods. The modulation selection rule of our method is independent of the channel statistics. We compare the performance of our method with several other methods that are optimized for the assumed channel statistics. Finally, we address the practical consideration of the AM such as CSI mismatch, co-channel interference, outage probability, and adaptation rate.

In chapter 4, we extend the long-range prediction algorithm into frequency domain. We derive the theoretical linear MMSE filter and discuss the factors that affect the prediction accuracy. Next, we employ the least-square (LS) methods in conjunction with forward and backward linear prediction to estimate the correlation matrix and obtain the filter coefficients. Robust adaptive LRP techniques based on the least-mean-squares (LMS) algorithm that can track the time-variant channel statistics are developed as well. We analyze the transient behavior, convergence rate and asymptotic mean-square error of the LMS algorithms. Lastly, we employ the Jakes and the physical model to test our proposed system and demonstrate the significant gains relative to non-adaptive techniques for sufficiently correlated channels and prediction range. In addition, typical and challenging cases are identified to test the limitation of the adaptation rate of our system.

In chapter 5, the multicarrier OFDM system is introduced. We first revive the modulation and demodulation process of the OFDM system and analyze the effect of Doppler shifts on the receiver SNR performance. Then we apply the long-range channel prediction and adaptive modulation on the OFDM system. In particular, the desired carrier CSI is predicted based on the observation of multiple subcarriers. Modified channel prediction method for OFDM is proposed to significantly reduce the complexity based on the fact that all the subcarriers have approximately the same Doppler shift associated with each path. Adaptive LRP methods based on the LMS and RLS algorithms that do not require the knowledge of channel statistics are developed. Next, optimal and less complex sub-optimal bit and power allocation algorithms are used and the performances are compared for the imperfect CSI. Subsequently, reduced feedback methods are developed to reduce the overhead to synchronize the CSI at the transmitter and the receiver. Finally, the modified

Jakes and physical models are used to test the performance of the adaptive OFDM system aided by the long range prediction.

CHAPTER 2

FADING CHANNEL CHARACTERIZATION AND MODELING

2.1 Background

Radio-wave propagation through wireless channel is a complicated phenomenon characterized by many factors such as the reflection, diffraction, and scattering. It is too complex to precisely describe this phenomenon by a mathematical model. However, considerable efforts and extensive measurements have been made using the continuous−wave or narrowband signals to determine the propagation characteristics. The measured channel is analyzed by data reduction and characterized by statistical properties that are used by system designers. The results are a range of relatively simple and accurate statistical models for fading channels, which depend on the particular propagation environment and the underlying communication scenario.

As a mobile moves over a very short distance, the instantaneous received signal strength may fluctuate rapidly resulting in *small-scale fading* [51]. (Propagation models that describe the mean signal strength for a large transmitter-receiver separation distance are called *large-scale fading* models.) The reason for this is that the received signal is a sum of may contributions coming from different directions. Throughout the thesis, we only consider the *small-scale fading*. Interested readers are referred to textbooks, e.g. [6] and [51], for different types of fading channels.

In this chapter, we first briefly review a commonly used continuous-time mobile fading channel model, which can be described as a weighted sum of complex sinusoids. This model is based on the assumptions of planar wave-front, constant vehicle velocity and

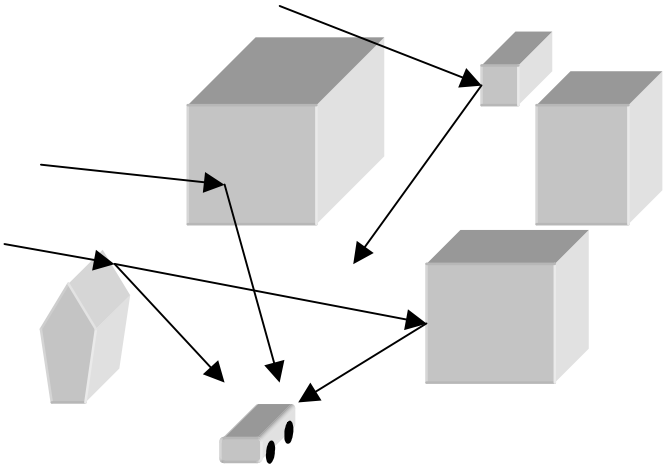


Figure 2.1 Scattering from objects results in signals arriving at the mobile station from many directions.

propagation via reflecting and scattering objects. The assumptions and approximations in the derivation of this model will be discussed. Channel correlation functions over time and frequency based on this model will be established. The limitation of this simple sum-of-sinusoids model is discussed, and a non-stationary physical model that can generate realistic mobile fading channels [16,33,34] is introduced to test our system performance in the later chapters.

2.2 Frequency Non-Selective Multipath Fading Channels

A typical scenario of a mobile station (MS) traveling along a street is depicted in Figure 2.1, where we have shown a few of the many possible rays arriving at the MS. These paths, which make it possible to communicate between base stations (BS) and the MS without line-of-sight (LOS), involve physical phenomena of reflection, diffraction, and scattering from buildings, vehicles, people, and so on. The envelope and phase of the received signal fluctuate over time due to the superposition of the multiple waves leading to either constructive (peak) or destructive (deep fades) interference as shown in Figure 2.2.

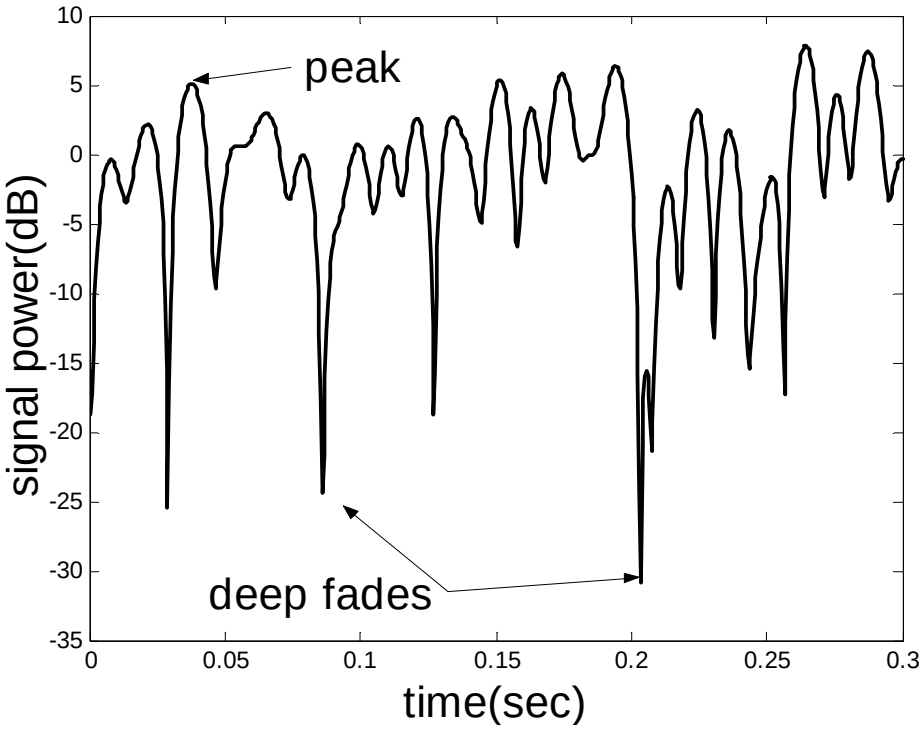


Figure 2.2 Example of signal fading over time.

This fading phenomenon, which is called *multiplath fading*, can severely degrade system performances unless measures are taken to compensate at the receiver. In chapter 3, we will employ adaptive modulation and power control techniques at the transmitter based on the feedback channel information to mitigate the performance loss due to the violent channel variations. Therefore, it is important to first realize the properties of the multipath-fading channels.

2.2.1 Mathematical Model

As shown in Figure 2.1, many multiple plane waves arrive at a MS from many different directions. Associated with each path is a propagation delay and attenuation factor. Due to changes in the structure of the medium, both the propagation delay and the

attenuation factors are time-variant. The channel impulse response $h(t, \tau)$ at time t due to an impulse applied at time $t - \tau$ can be expressed as:

$$h(t, \tau) = \sum_{n=1}^{N} A_n(t)\delta(\tau - \tau_n(t)) \tag{2.1}$$

where $\delta(\bullet)$ is the Dirac delta function, N is the number of waves, and for the n^{th} path, $A_n(t)$ is the (real) attenuation factor and $\tau_n(t)$ is the propagation delay. By taking the Fourier transform of $h(t, \tau)$ with respect to τ, we obtain the time-variant frequency response:

$$H(t, f) = \int_{-\infty}^{\infty} h(t, \tau)\exp\{-j2\pi f\tau\}d\tau = \sum_{n=1}^{N} A_n(t)\exp\{-j2\pi f\, \tau_n(t)\} \tag{2.2}$$

We now consider a narrowband signal

$$s(t) = \mathrm{Re}\{s_l(t)exp(j2\pi f_c t)\} \tag{2.3}$$

where $s_l(t)$ is the complex low-pass signal [45] and f_c is the carrier frequency. Assume the symbol interval is much larger than the excessive propagation delay and hence the channel is *frequency non-selective* [45]. This implies that all of the frequency components in s(t), denoted as S(f), undergo about the same attenuation and phase shift. Since S(f) has its frequency content concentrated in the vicinity of f_c, the received complex lowpass signal $r_l(t)$ over the time-variant channel (2.2) can be approximately expressed as [45]:

$$r_l(t) \approx \int_{-\infty}^{\infty} H(t, f_c)\, S_l(f)exp\{j2\pi ft\}df = H(t, f_c)\int_{-\infty}^{\infty} S_l(f)exp\{j2\pi ft\}df =$$

$$H(t, f_c)\, s_l(t) = s_l(t)\sum_{n=1}^{N} A_n(t)\exp\{-j\phi_n(t)\} \tag{2.4}$$

where $S_l(f)$ is the frequency response of $s_l(t)$ and $\phi_n(t) = 2\pi f_c\tau_n(t)$. From (2.4), the received signal $r_l(t)$ is the transmitted signal $s_l(t)$ multiplied by a time-variant complex channel gain

H(t, f_c), which is called *channel state information* (CSI). As indicated in (2.4), CSI consists of a sum of a number of time-variant vectors having amplitudes $A_n(t)$ and phases $\phi_n(t)$. For carrier frequency f_c in the ultrahigh frequency (UHF; 300MHz ≤ f ≤ 3GHz) and delay propagation $\tau_n(t)$ on the order of 0.1 to 10μsec. It is reasonable to model $\phi_n(t)$ using a uniform distribution around $[0,2\pi]$. For large N, the CSI is distributed approximately as a zero mean complex Gaussian random variable by central limit theorem [43]. Hence the amplitude $\alpha(t) = |r_l(t)|$ is *Rayleigh* distributed with probability density function (*pdf*):

$$p(\alpha) = \frac{2\alpha}{\Omega} exp\{\frac{-\alpha^2}{\Omega}\} \tag{2.5}$$

where $\Omega = E[\alpha^2]$ is the average power of the CSI. When there is a dominant (non-fading) signal component present, such as LOS propagation path, the envelope distribution is *Ricean* [51]. Throughout the thesis, we analyze our system performance over the *Rayleigh* fading channel.

2.2.2 Sum-of-Sinusoids Model

The CSI modeled in (2.4) does not give much information about the mechanism of its time-variant properties. In this section, we show the variation of the multipath fading channel significantly depends on the velocity and the direction of the MS, and we establish a *sum-of-sinusoids* (SoS) model. As shown in Figure 2.3, consider a mobile moving at a constant velocity *v* from X at $t = t_0$ to Y in a very short period of time Δt while it receives signals from a remote source S. The distance between X and Y is *v*Δt. The incident angle θ can be considered the same at X and Y since the source S is assumed to be very far away. Let $\tau(t_0)$ and $\tau(t_0 + \Delta t)$ denote the propagation delay from S to point X and Y, respectively. Then $\tau(t_0 + \Delta t)$ can be approximately calculated as:

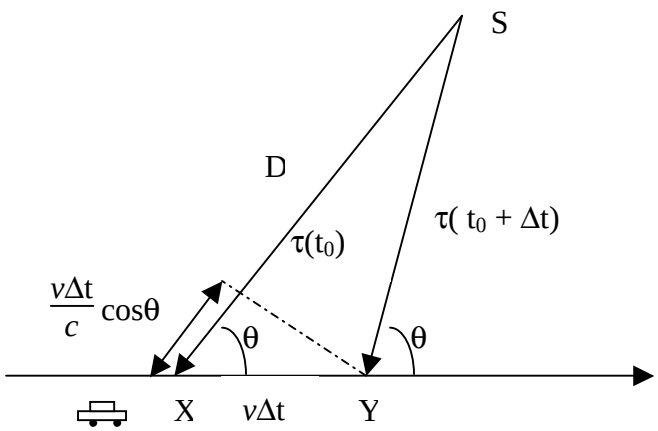


Figure 2.3 Illustration of short-term channel behavior

$$\tau(t_0 + \Delta t) \approx \tau(t_0) - \frac{v\Delta t}{c}\cos\theta \tag{2.6}$$

where c is speed of the light. By using the approximation (2.6), (2.2) can be expressed as:

$$H(t_0 + \Delta t, f_c) = \sum_{n=1}^{N} A_n(t_0 + \Delta t) exp\{-j2\pi f_c \tau_n(t_0 + \Delta t)\}$$

$$\approx \sum_{n=1}^{N} A_n(t_0 + \Delta t) exp\{j(2\pi f_d(n)\,\Delta t - \phi_n)\} \tag{2.7}$$

where $\phi_n = 2\pi f_c \tau_n(t_0)$, $f_d(n) = f_c \frac{v}{c}\cos\theta_n$ is called Doppler shift, which accounts for the frequency shift due to the movement of the mobile relative to the signal source, and $f_c \frac{v}{c}$, denoted as f_{dm} is defined to be the maximum Doppler shift. In general, it requires large dynamic variations in the medium to cause significant changes for $A_n(t)$ [45]. Therefore, in a short period of time Δt, $A_n(t)$ can be modeled as a constant. Therefore, from (2.7), the channel variation over time Δt relative to instance t_0 at carrier frequency f_c can be modeled as a *sum-of-sinusoids* process:

$$c(t) = \sum_{n=1}^{N} A(n) \exp\{j(2\pi f_d(n)\, t - \phi(n))\} \tag{2.8}$$

It is observed that the Doppler shifts determine how rapidly the channel fluctuates over time. In [45], the *coherence time* $(\Delta t)_c$, a time period within which the CSI are strongly correlated, is defined as:

$$(\Delta t)_c \approx \frac{1}{2f_{dm}} \tag{2.9}$$

Thus, the CSI fades about $(\Delta t)_c$ sec away. In order to have a uniform measure, we multiply (2.9) by the maximum Doppler shift and the coherence time is expressed in terms of spatial movements of the antenna in terms of wavelength ($f_{dm}(\Delta t)_c = \frac{v(\Delta t)_c}{\lambda} = 0.5$). In other words, fades occur at least half a wavelength apart.

It is also interesting to investigate the relationship between CSI at two carrier frequencies. Let $f^2 = f_c + \Delta f$ be a carrier frequency Δf away from f_c. From (2.8), the CSI $c(f^2,t)$ at carrier f^2 can be expressed as [35]:

$$c(f^2,t) = \sum_{n=1}^{N} A(n)\exp\{j(2\pi f_d(n)'t - \phi'(n))\} \tag{2.10}$$

where $f_d(n)' = f_d(n) + \frac{\Delta f\, v}{c}$, and $\phi'_n = \phi_n + 2\pi\Delta f\tau_n(t_0)$. In general, Δf, for our interests, is on the order of hundred kilohertz, and thus $\frac{\Delta f\, v}{c}$ is very small and can be neglected. Therefore, $f_d(n)' \approx f_d(n)$ and it is the phase difference $2\pi\Delta f\tau_n(t_0)$ that dominates the correlation between the two carrier frequencies.

2.3 Channel Statistics and Measurement

We shall now develop a number of useful statistical properties that will be employed throughout the thesis based on the SoS model to characterize the multipath fading channel.

Let $\{\underline{A},\underline{\theta},\underline{\tau},\underline{\phi}_1\}$ denote the set $\{A(n),\theta(n),\tau(n),\phi_1(n), n=1\ldots N\}$ that parameterizes the channel model in (2.8) and (2.10).

2.3.1 Channel Correlation Functions

We develop two useful correlation functions that will be employed for the analysis of our prediction algorithm. For a deterministic channel where $\{\underline{A},\underline{\theta},\underline{\tau}\}$, are known, the *temporal average correlation function* (TACF) [28] can be calculated from (2.8) and (2.10) as:

$$R_T(\Delta t, \Delta f) = \lim_{T\to\infty} \frac{1}{2T-|\Delta t|} \int_{-T+|\Delta t|/2}^{T-|\Delta t|/2} c(f^1,t-\Delta t/2)\, c^*(f^2,t+\Delta t/2)dt =$$

$$\sum_{n=1}^{N} A(n)^2 exp\{-j2\pi\, f_d(n)\Delta t\} exp\{j2\pi\Delta f\tau(n)\} \qquad (2.11)$$

Note that TACF does not depend on the parameters $\{\underline{\phi}_1\}$. Assume $\{\underline{\phi}_1\}$ are mutually independent random variables uniformly distributed around $[0,2\pi]$. It can be readily shown from (2.8) and (2.10) that the conditional expectation $E_{\{\underline{\phi}_1\}}[c(f^1,t)c^*(f^2,t+\Delta t)|\{\underline{A},\underline{\theta},\underline{\tau}\}]$ = TAFC in (2.11), where $E_{\{\underline{\phi}_1\}}$ denotes the statistical expectation over all possible random parameters $\{\underline{\phi}_1\}$.

For c(t) characterized as wide sense stationary uncorrelated scattering (WSSUS) [45], the *ensemble average correlation function* (EACF) for two fading signals with the time difference Δt and the frequency separation Δf is defined as $R_E(\Delta t ,\Delta f) = E[c(f^1,t)\, c^*(f^2,t+\Delta t)]$, where E denotes the statistical average over all possible random parameters $\{\underline{A},\underline{\theta},\underline{\tau},\underline{\phi}_1\}$. Equivalently, $R_E(\Delta t ,\Delta f)$ is the average of the TACF, i.e., $R_E(\Delta t ,\Delta f) = E_{\{\underline{A},\underline{\theta},\underline{\tau}\}}\{E_{\{\underline{\phi}_1\}}[c(f^1,t)c^*(f^2,t+\Delta t)|\{\underline{A},\underline{\theta},\underline{\tau}\}]\}$. It is shown in the Appendix A the $R_E(\Delta t ,\Delta f)$ can be

factored into the time-domain correlation function $R_t(\Delta t)$ and the frequency domain correlation function $R_f(\Delta f)$ as

$$R_E(\Delta t, \Delta f) = \Omega R_t(\Delta t) R_f(\Delta f) \tag{2.12}$$

where $\Omega = E[|c(f^i, t)|^2]$ is the average power of the fading signals. (We normalize Ω to 1 throughout the paper.) Assume $\theta(n)$ is uniformly distributed around 2π (*isotropic scattering* [54]), and the propagation delay $\tau(n)$ is exponentially distributed [35] with the probability density function

$$p(\tau) = \frac{1}{\sigma} exp\{-\tau/\sigma\} \tag{2.13}$$

where σ is the *rms* delay spread [51]. Then

$$R_t(\Delta t) = J_0(2\pi f_{dm}\Delta t) \tag{2.14}$$

is the zero order Bessel function [35] and $R_f(\Delta f) = \frac{1}{1+(2\pi\Delta f\sigma)^2} + j\frac{2\pi\Delta f\sigma}{1+(2\pi\Delta f\sigma)^2}$. We define $f_{dm}\Delta t$ and $\Delta f\sigma$ as the *normalized time difference* (NTD) and the *normalized frequency separation* (NFS), respectively. Note that (2.12) is expressed in terms of these normalized quantities. The power spectrum density $S_c(f)$ is obtained by taking the Fourier transform of $R_t(\Delta t)$ as [35]:

$$S_c(f) = \begin{cases} \frac{\Omega}{2\pi fdm} \frac{1}{\sqrt{1-(f / fdm)^2}} & |f| \le fdm \\ 0 & \text{otherwise} \end{cases} \tag{2.15}$$

Note the power spectrum is zero outside the maximum Doppler shift. This property can be employed for noise reduction by suppressing out-of-band signals [19].

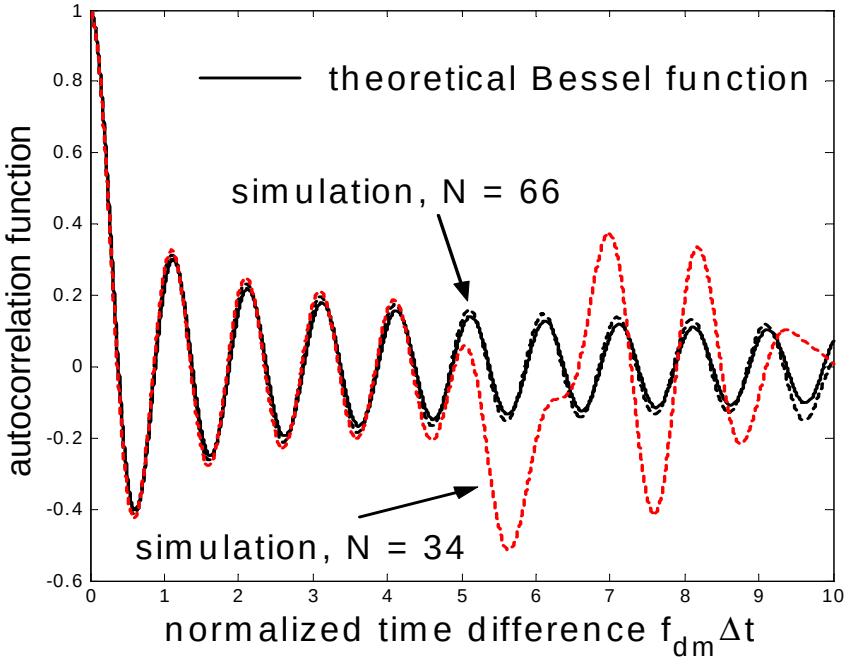


Figure 2.4 Comparison of the theoretical correlation function and the numerical results from Jakes model.

2.4 Laboratory Simulation

To generate the mobile fading channel for system simulations, the most accurate method is to use recoded strips of actual channel measurements. Unfortunately, it is time-consuming and often inconclusive due to uncertainty in the statistical signal variations actually encountered. Hence a simple channel simulator that can generate the assumed statistical properties is desired. One technique of this type has been suggested by Jakes in 1974 [35].

The description of the Jakes method begins with (2.8) and assumes N equal strength multipath components uniformly distributed in angle, i.e.,

$$\theta_n = \frac{2\pi n}{N}, \qquad n = 1,2,\ldots,N \tag{2.16}$$

While in theory very large N is required for the TACF $R_T(\Delta t, 0)$ (2.11) of a deterministic channel process to approach $R_t(\Delta t)$ in (2.14), it is shown by Jake that the time domain correlation function $R_t(\Delta t)$ can be well approximately by $R_T(\Delta t, 0)$ for a summation of a relative small number N of sinusoids [35]. Detailed analysis of the Jakes model can be found in [35,48]. Figure 2.4 shows the comparison of the theoretical correlation (Bessel) function (2.14) and the numerical results of $R_T(\Delta t, 0)$ for different number of path N from the Jakes model. It is observed that larger N is required for larger normalized time difference $f_{dm}\Delta t$ for accurate modeling of the Rayleigh fading channel using the Jakes model.

2.4.1 Modified Jakes Model: Random Phase Model

By designating the fixed incident angles $\theta(n)$, Jakes model can easily generate the desired time domain correlation function. However, in the real world, $\theta(n)$ is randomly distributed on $[0,2\pi]$, and in both LOS and non-LOS environment, spectral analysis of measured fading data strongly supports the conjecture that the complex baseband fading process mainly consists of a small number of sinusoids. In addition, for our adaptive channel prediction method, which is capable of tracking the current channel statistics, the behavior (learning curve, prediction accuracy) of the predictor is mainly determined by the TACF, not the EACF. Therefore, instead of using $\theta(n) = 2\pi n/N$, we generate the fading channel using random uniform distribution of the incident angle $\theta(n)$ on $[0,2\pi]$. We call this method *random phase model* (RPM). Unlike the Jakes model where the $R_T(\Delta t, 0)$ is fixed for each generated channel, RPM can generate different $R_T(\Delta t, 0)$ to investigate the performance of our system in different channel environments. After generating the CSI at frequency f^i, we compose the CSI at frequency $f^j = f^i + \Delta f$ by employing the same parameters $\{\underline{A},\underline{\theta},\underline{\tau}\}$ except

phases $\phi_i(n) = \phi_j(n) - 2\pi\Delta f\tau(n)$, n = 1…N, where $\{\underline{\tau}\}$ is exponentially distributed according to (2.13).

2.5 Limitation of Sum-of-Sinusoids Model and Physical Model

The Jakes model and RPM can easily generate the multipath Rayleigh fading signals with desired correlation functions $R_T(\Delta t, \Delta f)$ (2.11) by assuming amplitudes, phases, and Doppler shifts are time-invariant. However, in a real mobile radio channel, the shape of the correlation functions, affected by many time-variant factors such as the vehicle speed, number and location of the scatters, also vary with time. These models are sufficient for systems concerned with short term and narrowband channel behavior. However, in our study of long-range channel prediction, we predict the CSI over time and frequency domain, and the observation interval and the memory span are much larger. This requires us to take into account the variation of parameters such as the excess delays, the Doppler shifts, and the amplitudes. Thus, realistic non-stationary modeling is necessary.

A novel physical model based upon the method of images combined with diffraction theory is proposed in [16,33,34]. This model can (1) create non-stationary datasets to test our long range prediction algorithm, (2) classify the reflector geometries that have typical or the most severe parameter variations, so that the reflector configurations for test datasets can be appropriately chosen, (3) provide limits on the speed of adaptation needed for an algorithm to predict the channel significantly into the future, and thereby reveal the timing of future deep fades, and (4) illuminate the origins of the temporal and statistical properties of the measured data [34]. It is demonstrated in [16,33,34] that this physical model generates datasets that closely resemble measured data, and results of the long range prediction for the physical model and measured data are similar, and differ significantly from those produced for the

Jakes model. The generation and the properties of the deterministic physical modeling have been addressed in [16,33,34]. We next briefly review these properties.

The first property is that the parameters A(n), f_d(n),ϕ(n), and τ(n) vary much slower than the actual fading channel. The rate of change depends on the local environment. Insights into this property can be found in [33,34]. The variation of these parameters limits the range for which the fading process can be reliably predicted in time and in frequency given current and past fading channel observations. The second important property is that the number of significant scatterers N is modest. There are likely to be many insignificant scatterers that add small variations to the received signals. However, we only want to track those beams with power up to 1/10 the power of the strongest beams. Propagation studies show that the number of significant reflectors is modest, usually not more than 20 [35,50].

In chapters 4 and 5, we illustrate the geometries for generating the physical model to test our prediction algorithm and evaluate the performance loss relative to the standard sum-of-sinusoids model due to the time-variant non-stationary processes.

CHAPTER 3

ADAPTIVE MODULATION

3.1 Background

As we discussed in the chapter 1, adaptive transmission (AT) techniques are now widely recognized as a key solution to significant improvement of the wireless system performance in hostile fading channels. The basic idea of AT techniques is to adapt the transmission parameters to take advantage of prevailing channel conditions. A simple AT system diagram is shown in Figure 3.1. The Transmitter sends data symbols through wireless mobile channel *A* (narrow band or wideband) and obtains estimated CSI (or system parameters) from the receiver via channel B for adjusting the transmission modes to optimize the system performance. If the Time Division Duplex (TDD) scheme [57] is employed, the channel A and B share the same frequency band but different time slots to provide both a forward and reverse link. In this case, the transmitter can estimate the CSI from the received signals if the channel is considered *reciprocal*, which means that the CSI are strongly correlated in a forward and reverse time frame. This scheme is called *open-loop* system. On the other hand, if the Frequency Division Duplex (FDD) is used, then the CSI is estimated at the receiver and fed back to the transmitter via the *feedback channel B.* (Alternatively, instead of the estimated CSI, transmission mode can be decided by the receiver and fed back.) This scheme is called the *closed-loop* system. The TDD adaptive transmission has the advantage of reducing feedback overhead. However, it is very sensitive to timing and has only recently been used for indoor or small area wireless applications where the physical coverage distances are much smaller than the many kilometers used in conventional cellular

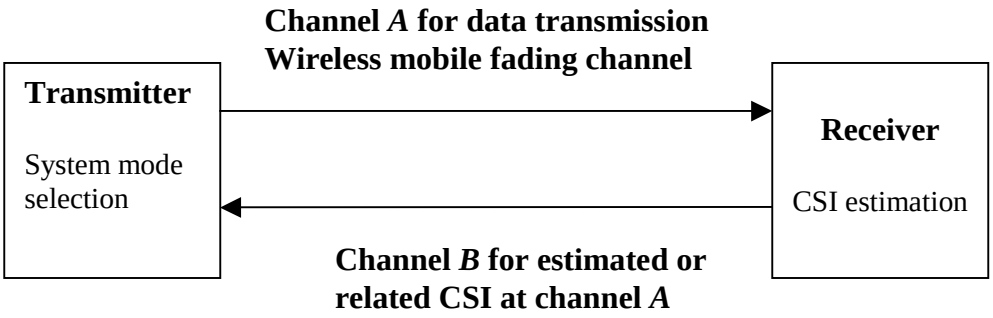


Figure 3.1 Adaptive transmission system diagram

systems [51]. Moreover, there is a time-latency due to the fact that communications is not full duplex. In this and the following chapters, we mainly focus on the *closed loop* adaptive modulation (AM) system in conjunction with power control (PC) to maximize the system capacity.

Having discussed briefly the principles of AM and the possible scenarios where it can be applied, we now consider the practical limitation and implementation issues. The following steps have to be considered to design an AM system:

- ***Channel quality estimation***: In order to appropriately select the system parameters to be employed for the next transmission, a reliable estimate of the CSI during the next active transmission time slot is necessary. Pilot symbol assisted modulation (PSAM) [9,55] has been proposed as an attractive technique to detect the CSI in the rapid fading environment by periodically inserting known symbols, from which the receiver derives its amplitude and phase reference. In addition, channel prediction techniques [3,16,19,20,27] can be used to mitigate the mismatch between the estimated CSI and the actual CSI experienced for the next transmission symbol.
- ***Determination of the appropriate parameters:*** Based on the estimated or predicted channel conditions, the transmitter has to select an appropriate transmission mode.

Historically, interests in techniques of adapting the modulation and transmission rate parameters began in 1968, when Hayes [25] adapted the signal amplitude according to the prevalent channel environment by utilizing a feedback channel between the transmitter and the receiver. More recently, this area has attracted enormous research interests due to the large demands of voice and multimedia service and the significantly enhanced hardware technologies. Most of this research work focuses on the following issues: modulation selection rule [14,21,24,49,57,64], transmit signal power control [7,21,37,56], symbol rate switching [38,57], adaptive coding techniques [22,24,42] and AM in conjunction with antenna diversity [30,32,62,65].

- ***Adaptation rate***: The adaptation rate determines what kind of channel variations the AM algorithm is tracking. If the channel is changing faster than it can be estimated and fed back to the transmitter, AT techniques will perform poorly, and other means to mitigate the effects of fading should be used. It is easy to understand that faster adaptation leads to larger capacity gain, since the channel variations are exploited in a more accurate manner. However, fast adaptation has practical limitations such as hardware constraints. Besides, fast adaptation increases the number of mode-change messages sent to the receiver, which consume bandwidth and time resources.
- ***Feedback:*** The feedback messages inform the transmitter of the CSI estimated or the transmission mode decided by the receiver. The feedback load should be minimized since it consumes resources that would be otherwise used for data. For example, channel mean feedback instead of the exact CSI is employed in [66] to reduce the feedback load. In chapter 5, we develop several reduced feedback methods for the mobile radio adaptive OFDM system.

- ***Signaling or blind detection of the employed parameters***: The transmitter has to inform the receiver of the transmission mode. This information can either be conveyed within the symbol, at the cost of loss of effective data throughput, or the receiver can attempt to estimate the parameters by means of blind detection techniques [60].

In this chapter, adaptive modulation employing Multilevel Quadrature Amplitude Modulation (MQAM) in conjunction with power control (PC) is applied in a narrow-band Rayleigh fading channel to maximize the transmission throughput. AM for the wideband multicarrier OFDM system is addressed in the chapter 5.

3.2 Fading Channel Capacity

The ideal system achieves a small bit error rate with a high spectral efficiency and a low SNR. Information theory [13] places bounds on the value of these parameters that can be achieved by any coding and modulation techniques. The capacity of the Rayleigh fading channel is addressed in [23]. We first briefly review the fading channel capacity to provide invaluable design concepts, and the results serve as the theoretical performance upper bound.

Consider a discrete-time narrow-band system model depicted in Figure 3.2 with a *stationary* and *ergodic* complex fading channel gain $c[i]$ with average power $\Omega = E[|c[i]|^2]$ and additive white Gaussian noise (AWGN) $n[i]$ with power spectrum N_0. For an average transmitted power constraint $\bar{S}$, the average received SNR $\bar{\gamma} = \bar{S}\Omega/N_0$, and the instantaneous received SNR $\gamma[i] = \bar{S}\alpha^2[i]/N_0$, where $\alpha[i] = |c[i]|$ is the amplitude of the complex channel gain. We normalize the average channel power $\Omega = E[\alpha^2] = 1$ and thus $\bar{\gamma} = \bar{S}/N0$. Suppose power control is employed and denote $S[i]$ as the transmitted power at time i based on the parameter $\gamma[i]$. (Note $E[S[i]] \leq \bar{S}$.) We assume *coherent detection* and the phase variations are

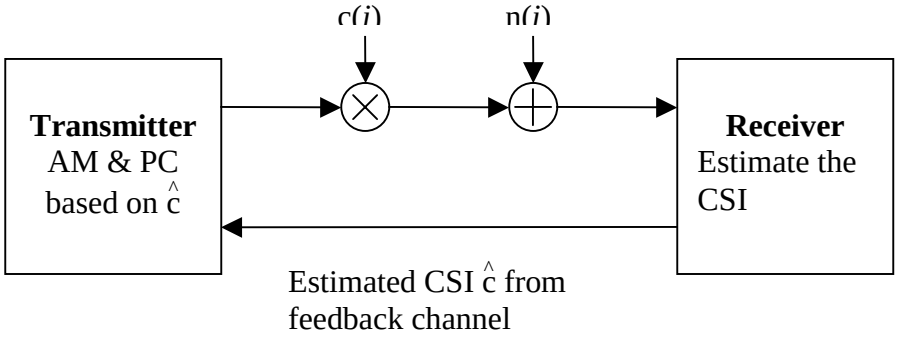


Figure 3.2 System model of adaptive modulation and power control

completely removed during the demodulation process. Therefore, the receiver BER performance is decided by the received SNR $\gamma_S = \gamma[i]S[i]/\bar{S}$. Since $\alpha[i]$ is stationary, we will drop the time index *i* for simplicity. We assume α is Rayleigh fading (2.5) distributed in the numerical simulation below, although our formulas apply for any distribution of α.

From [15], the GAP concept can be used to approximate the performance of a coding and modulation (C/M) scheme in use on an AWGN channel. For a gap Γ, corresponding to the target probability of error and the coding scheme in use, the number of bits *b* that can be transmitted on a given AWGN channel using QAM as a function of received SNR γ_S is given as:

$$b = \log_2(1 + \frac{\gamma_S}{\Gamma}) = \log_2(1 + \frac{\gamma[i]S[i]}{\bar{S}\Gamma}) \qquad (3.1)$$

The parameter Γ denotes the maximum possible improvement gain for that C/M scheme relative to the channel capacity. For example, $\Gamma \approx 3$ for uncoded QAM with BER = 10^{-3}. The maximum throughput R_{max} over fading channel with GAP Γ and bandwidth B can be calculated as:

$$R_{max} = \max_{S(\gamma):\int S(\gamma)p(\gamma)d\gamma = \bar{S}} \int_0^{\infty} B\log_2(1 + \frac{\gamma S(\gamma)}{\bar{S}\Gamma})p(\gamma)d\gamma \quad (3.2)$$

Note that $\Gamma = 1$ in (3.2) results in the fading channel capacity C, i.e.,

$$C = \max_{S(\gamma):\int S(\gamma)p(\gamma)d\gamma = \bar{S}} \int_0^{\infty} B\log_2(1 + \frac{\gamma S(\gamma)}{\bar{S}})p(\gamma)d\gamma \quad (3.3)$$

For the Rayleigh fading channel, $p(\gamma)$ can be expressed as [58]:

$$p(\gamma) = \frac{1}{\bar{\gamma}} exp\{-\frac{\gamma}{\bar{\gamma}}\}, \gamma \geq 0 \quad (3.4)$$

where $\bar{\gamma} = E[\gamma] = \Omega\bar{S}/N_0$. From [23], the power control method to maximize (3.2) is

$$\frac{S(\gamma)}{\bar{S}} = \begin{cases} \frac{1}{\gamma_0} - \frac{1}{\gamma\Gamma} & \gamma \geq \gamma_0/\Gamma \\ 0 & \gamma \leq \gamma_0/\Gamma \end{cases} \quad (3.5)$$

where γ_0 is obtained by solving

$$\int_{\gamma_0}^{\infty} (\frac{1}{\gamma_0} - \frac{1}{\gamma}) p(\gamma)d\gamma = 1 \quad (3.6)$$

From the bit rate gain (3.2) and power adaptation (3.5), it is observed that the power adaptation and spectral efficiency for MQAM technique has an effective power loss of Γ relative to the optimal transmission scheme. That is, there is a simple relation between the maximum spectral efficiency of a fading channel and the spectral efficiency of uncoded adaptive MQAM methods [21].

For a constant transmit power case, the capacity of (3.3) reduces to:

$$C = \int_0^{\infty} B\log_2(1+S(\gamma)\gamma)p(\gamma)d\gamma \quad (3.7)$$

Note that (3.7) is the Shannon capacity of the fading channel when the transmitter adapts to the channel variation using a constant-power variable-rate strategy. It is shown in [23] the

difference between (3.7) and (3.3) is a small fraction of a decibel for most types of fading. This negligible impact of power adaptation results from the fact the capacity (3.7) places no restriction on the input signal constellation. Thus, when the constellation size is restricted to discrete size, restricting the power has a greater impact on spectral efficiency [21]. Next we introduce our discrete rate adaptive modulation in conjunction with continuous power control method.

3.3 Discrete AM and Continuous PC

Shannon capacity (3.3) and the continuous MQAM technique (3.1) places no restriction on the complexity of the multiplexed transmission scheme to achieve the optimal performance. In our study, we employ variable rate and variable power square multilevel quadrature amplitude modulation (M(i)–QAM, M(0) = 0, M(1)= 2, M(i) = $2^{2(i-1)}$, $i = 2 \ldots 4$) signal constellations as shown in Figure 3.3 due to their inherent spectral efficiency and ease of implementation [21,45]. Thus, at each symbol interval, we transmit a modulation level from the set {M(i): i = 0,1…4} based on the observed CSI. *Gray encoding* [45] is used to map the k information bits to the possible 2^k constellations. A Star–QAM adaptive modulation was studied in [64]. Let $2d$ be the Euclidean distance between the closest constellation points, then the average transmitted power $\bar{S}$ for BPSC, 4-QAM, 16-QAM, and 64-QAM in Figure.3.3 are d^2, $2d^2$, $10d^2$, and $42d^2$. The BER for each modulation on AWGN channel (channel gain α =1) can be represented by [12]:

$$P_e(\mathrm{M(i)}, \gamma) = \mathrm{Q}(\sqrt{2\gamma}), \qquad \mathrm{i} = 1$$

$$P_e(\mathrm{M(i)}, \gamma) = \frac{1}{\log_2\sqrt{\mathrm{M}}} \sum_{k=1}^{\log_2\sqrt{\mathrm{M}}} P_b(k), \qquad \mathrm{i} = 2 \ldots 4 \tag{3.8}$$

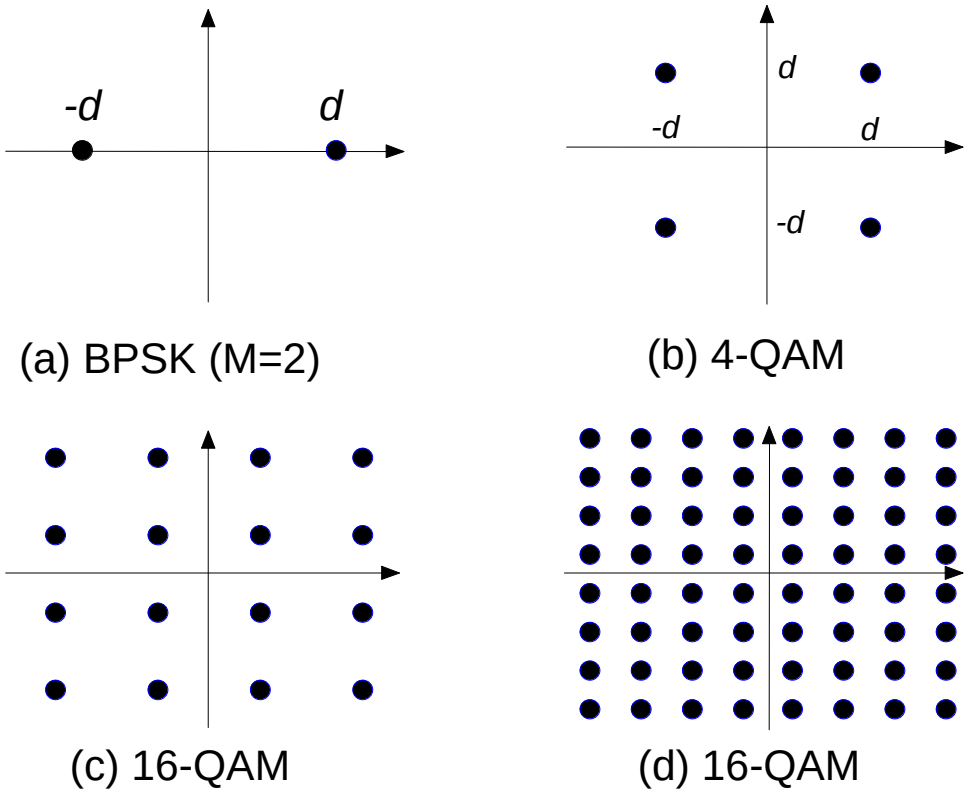


Figure 3.3 Square M-QAM constellation with Gray coding

where $\gamma = \bar{S}/N_0$ is the SNR per symbol and

$$P_b(k) = \frac{1}{\sqrt{M}} \sum_{i=1}^{(1-2^{-k})\sqrt{M}-1} \{(-1)^{\lfloor (i2k-1)/\sqrt{M} \rfloor} (2^{k-1} - \lfloor \frac{i2^{k-1}}{\sqrt{M}} + \frac{1}{2} \rfloor) \, erfc((2i+1) \sqrt{\frac{3\gamma}{2(M-1)}})\} \quad (3.9)$$

For simplicity, the BER for i >2 can be calculated using an upper bound [21]:

$$P_e(M(i), \gamma) \leq 0.2 exp\{\frac{-1.5\gamma}{M(i)-1}\} \quad (3.10)$$

Figure 3.4 shows the BER vs. average SNR per symbol γ for M-QAM. It is observed that the exact BER is approximately upper-bounded by (3.10) for higher SNR. For example, the approximation can only be used when the SNR is higher than 10dB for 64-QAM and than 6dB for 16-QAM.

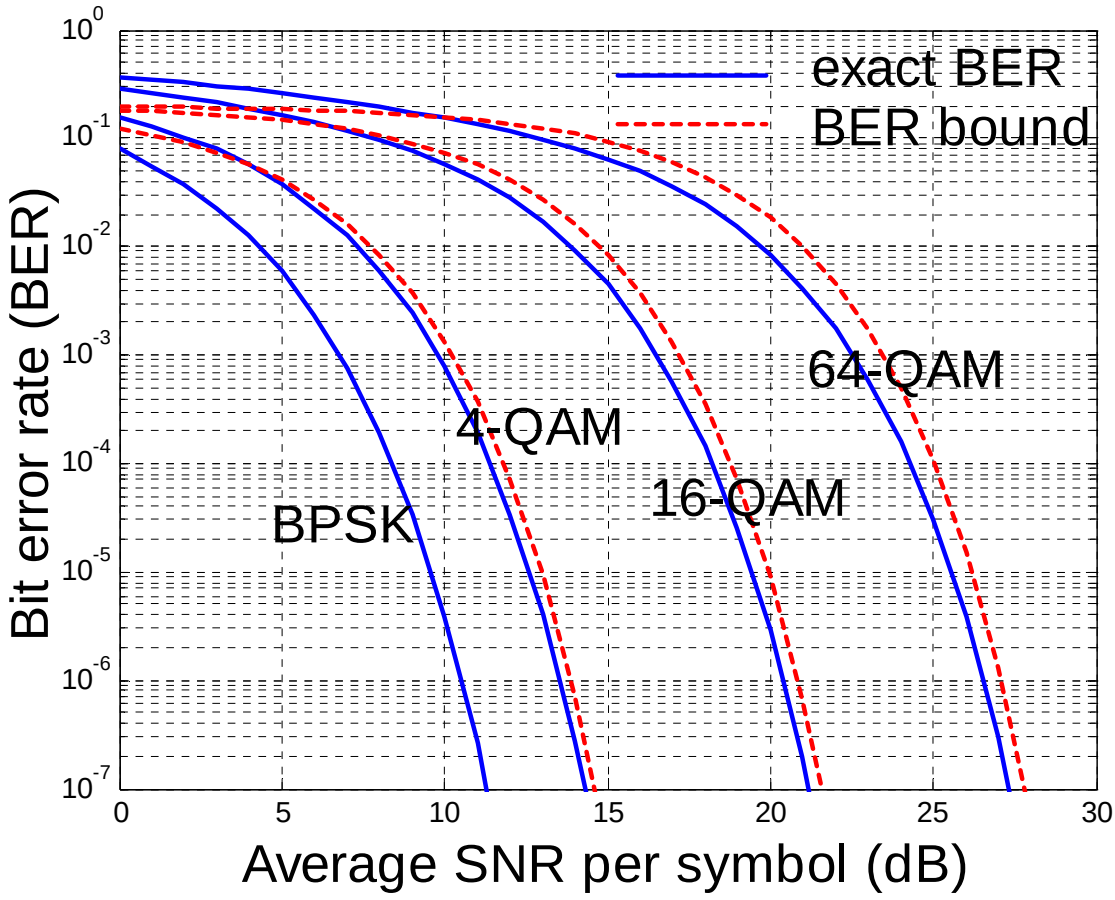


Figure 3.4 BER vs. average SNR per symbol for M-QAM

We now discuss our modulation selection and power adaptation algorithm. We first consider the case of perfect knowledge of CSI and instantaneous feedback. The practical issues of imperfect CSI are discussed in section 3.4. Given a target bit error rate BER_{tg} and a fixed transmit power per symbol $\bar{S}$ (or SNR $\bar{\gamma} = \bar{S}/N_0$), we adjust the modulation level according to the instantaneous channel gain α. We use threshold-based modulation selection method. The thresholds T_i, $i = 1 \ldots 4$ is the α that satisfies the condition $P_e(M(i), \bar{\gamma}\alpha^2) = BER_{tg}$, and when the channel gain α satisfies: $T_{i+1} \geq \alpha \geq T_i$, the modulation level $M = M(i)$ is employed, where $T_0=0$ and $T_5=\infty$. Figure 3.5 demonstrates an example of the threshold

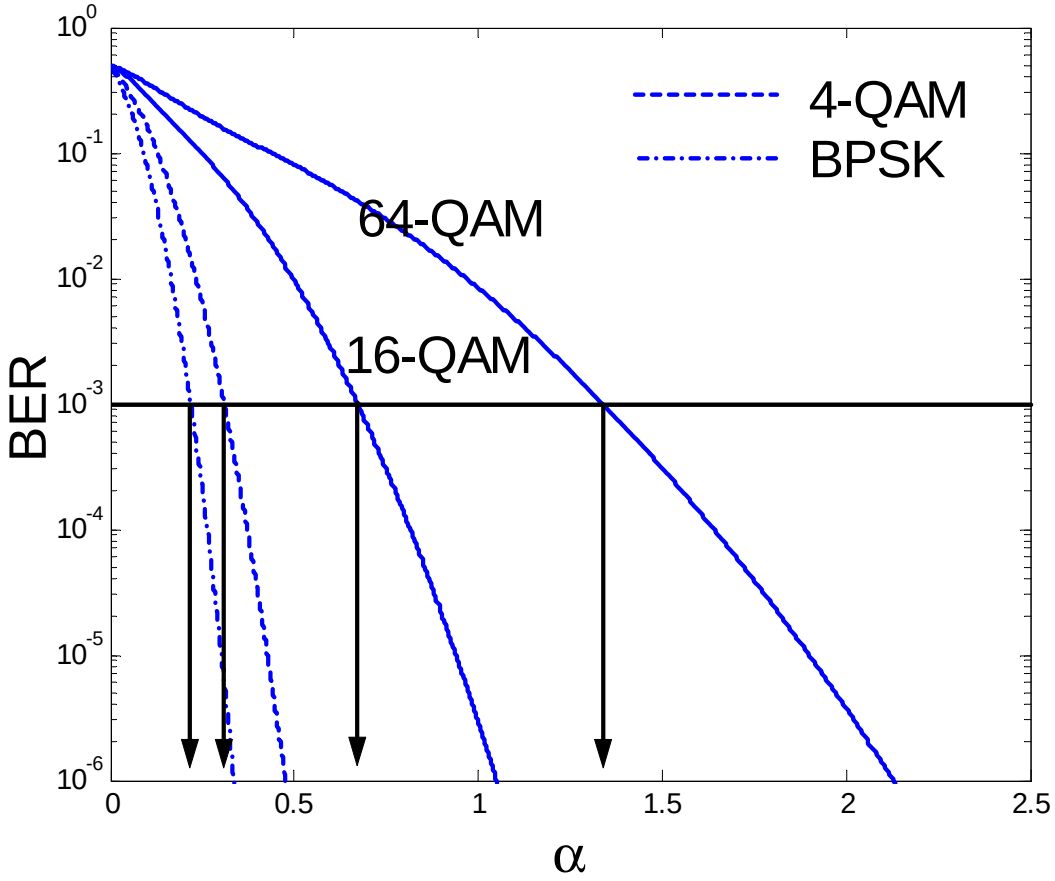


Figure 3.5 An example of modulation level-controlled adaptive modulation for perfect CSI. Average SNR = 20dB. BER = 10^{-3}.

search for $BER_{tg} = 10^{-3}$ and $\bar{\gamma}$ = 20dB. If the gap approximation (3.1) is used, thresholds can be simply determined by:

$$T_1 = \frac{Q^{-1}(BER_{tg})}{\sqrt{2\bar{\gamma}}}$$

$$T_i = \sqrt{\frac{(M(i)-1)\Gamma}{\bar{\gamma}}}, \qquad i=2\ldots4 \qquad (3.11)$$

where for $BER_{tg} = 10^{-3}$, $Q^{-1}(10^{-3}) \approx 3.0902$ and $\Gamma \approx 3$. We will use (3.11) to determine the thresholds for our numerical system simulation. Equivalently, the above modulation selection

algorithm is to choose the maximum modulation level such that the BER is less than the target BER for a given transmit SNR $\bar{\gamma}$ and channel gain α, i.e.,

$$M = \max\{\mathrm{M(i)} \mid P_e(\mathrm{M(i)}, \bar{\gamma}\alpha^2) < \mathrm{BER_{tg}}\} \tag{3.12}$$

Once the constellation M = M(i) and the threshold Ti are decided, the BER of this fixed power discrete rate adaptive modulation is lower than the BERtg since $P_e(M, \bar{\gamma}\alpha^2) < \mathrm{BER_{tg}}$ when α does not take on a threshold value. Therefore, we can use a power control policy to reduce the power consumption while maintaining the target BER. A fixed BER for the constellation M(i) can be maintained by multiplying $\bar{S}$ by a power control factor $\beta = (\mathrm{T_i}/\alpha)^2$ since

$$P_e(\mathrm{M(i)}, \bar{\gamma}(\frac{T_\mathrm{i}}{\alpha})^2\alpha^2) = P_e(\mathrm{M(i)}, \bar{\gamma}T_i^2) = \mathrm{BER_{tg}} \tag{3.13}$$

Therefore, the instantaneous transmitted power is $\bar{S}\beta$. Note the power control factor is always less than or equal to 1.

The modulation selection and power control algorithm described above is independent of the channel statistics. However, the average transmission rate and the average power usage depend on the probability density function of the channel gain α. The average bit rate is determine as:

$$\mathrm{R_{ada}} = \sum_{\mathrm{i}=1}^{4} \log_2 \mathrm{M_i} \int_{T_\mathrm{i}}^{T_{\mathrm{i}+1}} p_\alpha(\mathrm{x}) d\mathrm{x}, \tag{3.14}$$

This rate also gives the spectral efficiency assuming the ideal Nyquist data pulse [45]. And the average power usage can be calculated as:

$$\mathrm{S_{avg}} = \bar{\mathrm{S}} \sum_{\mathrm{i}=1}^{4} \int_{T_\mathrm{i}}^{T_{\mathrm{i}+1}} (\frac{T_\mathrm{i}}{\mathrm{x}})^2 p_\alpha(\mathrm{x}) d\mathrm{x} \tag{3.15}$$

The power control method above involves continuously varying transmitter power. It can be simplified by selecting a constant power for each constellation M(i) level to maintain the target BER. The power $E_s(i)$ for each constellation M(i) is chosen using:

$$\int_{T_i}^{T_{i+1}} P_e(\mathrm{M(i)}, \frac{E_s(i)\alpha^2}{N_0})\, p(\alpha \mid T_i < \alpha < T_{i+1}) d\alpha \leq \mathrm{BER_{tg}} \tag{3.16}$$

where $p(\alpha \mid T_i < \alpha < T_{i+1})$ is the conditional *pdf* of α is determined using thresholds T_i. Thus each modulation level is only associated with one transmit power. This is called discrete rate discrete power method [21].

3.3.1 Comparison of Different AM Methods

Several other adaptive modulation and power control algorithms are proposed in the literature. An optimal *discrete rate continuous power* scheme is proposed in [21], and an optimal *fixed power discrete rate* is derived in [14] by employing the *Lagrangian multiplier* technique to find a set of SNR-dependent AQAM mode switching level to maximize the achievable throughput while maintaining the average BER. While these methods are optimal, they require the knowledge of the channel statistics and complex numerical calculation to obtain the optimal threshold and thus are not practical. Our simpler *discrete rate continuous power* method results in less than 0.5 dB power loss compared with that in [21].

We compare several adaptive transmission techniques in Figure 3.6 over Rayleigh fading channel. Perfect CSI is assumed. A Continuous rate and power adaptation method in [21] is included in the comparison. Note that the method places no restrictions on the constellation size, which makes it impractical. We also found that the fixed-rate truncated channel inversion [21] based on M-QAM has similar performance to the discrete power discrete rate method, while it is also non-feasible in practice. We also plot the Shannon

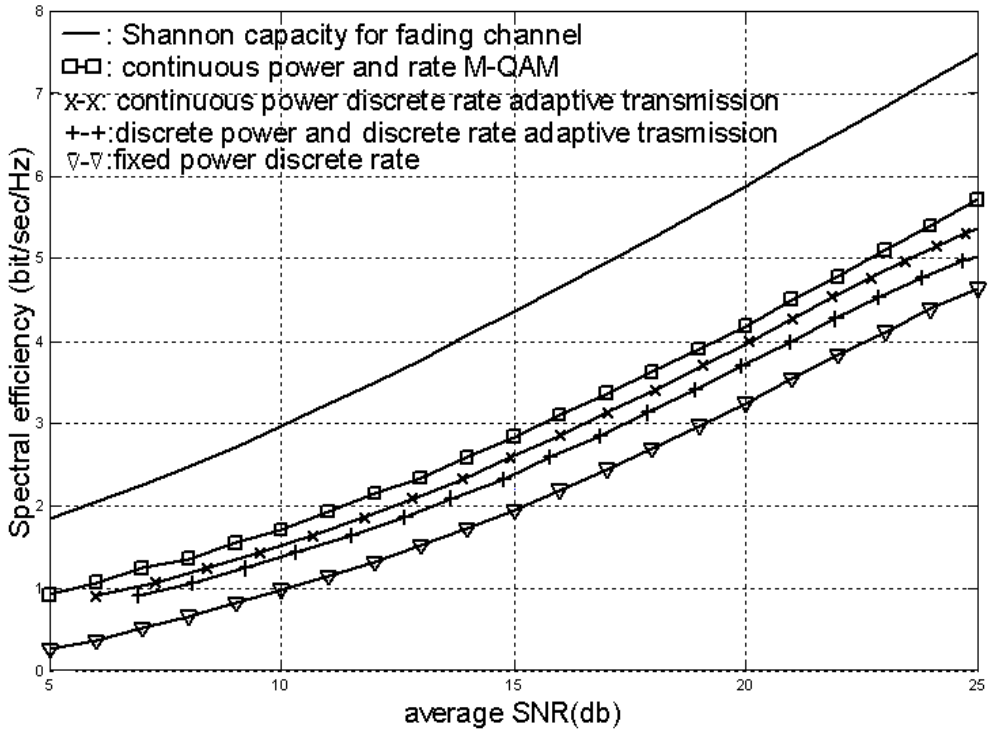


Figure 3.6 Spectral efficiency vs. SNR = S_{avg}/N_0 for different transmission techniques. Target BER = 10^{-3}.

capacity of the fading channel (3.3) for comparison. We observe that our continuous power control policy achieves about 3dB gain relative to the fixed power discrete rate adaptive modulation, and the discrete rate discrete power method has power loss of less than 2dB relative to the continuous power discrete rate transmission scheme.

3.3.2 Power Control Limitation

Cellular systems reuse the same frequency channel at spatially separated locations to provide more efficient utilization of the limited available spectrum within a given area [51]. Frequency reuse introduces co-channel interference, which ultimately determines the data rate and BER for each user. Thus, while power control techniques increases the power usage of a single channel, these techniques may also increase co-channel interference level in a

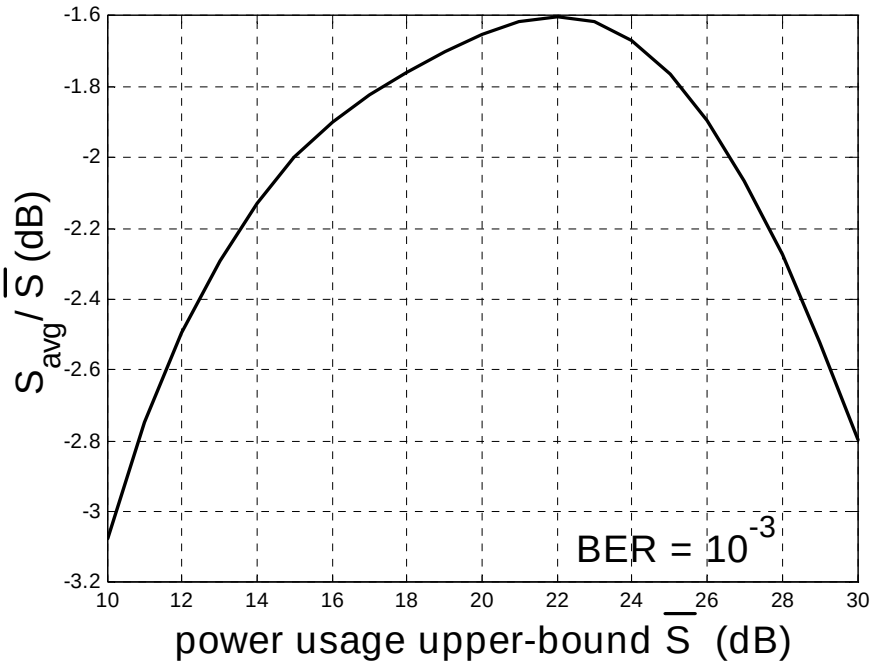


Figure 3.7 Comparison between the actual power usage and its upper bound.

cellular system, resulting in a higher reuses distance to mitigate the increased interference power.

For our power control method, the instantaneous transmitted power $\bar{S}\beta$ does not exceed $\bar{S}$ since the power control factor $\beta \leq 1$. Therefore, $\bar{S}$ is an important parameter in the design of a cellular system since it is the power usage upper bound during the transmission. Figure 3.7 shows the difference between the average power usage S_{avg} and power usage upper bound $\bar{S}$ over Rayleigh fading channel with target BER 10^{-3}. It is observed that there is about 1.5–3dB gap between the two parameters for our method depending on the upper bound $\bar{S}$. In general, a small gap between S_{avg} and $\bar{S}$ is desired. Larger gap may cause larger instantaneous power boost to increase CCI upon other users, which in turn may require further power increments for maintaining the target quality. In this case, although AM increases the spectral efficiency of a single channel, it may reduce the area spectral

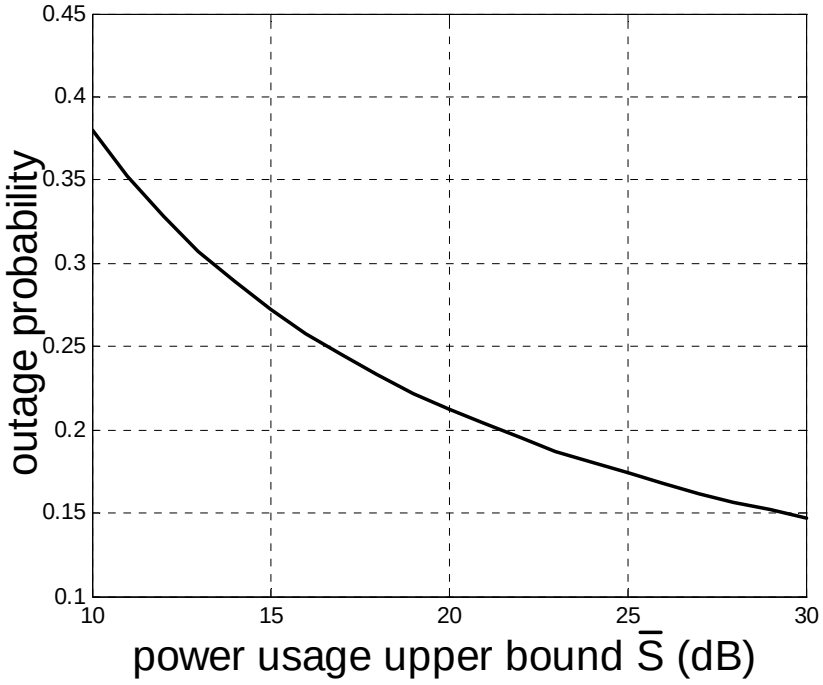


Figure 3.8 Outage probability vs. power usage upper bound $\bar{S}$. $BER_{tg} = 10^{-3}$.

efficiency. On the other hand, for the same upper bound $\bar{S}$, smaller gap can provide higher average power consumption and thus higher spectral efficiency.

3.3.3 Outage Probability

According to the modulation switching policy, the transmission is interrupted if the fading gain α falls below the threshold T_1. Thus, adaptive modulation suffers an outage probability P_{out} of:

$$P_{out} = \int_0^{T_1} p_\alpha(x)dx \tag{3.17}$$

Figure 3.8 shows the outage probability vs. the power usage upper bound $\bar{S}$ on the Raylegih fading channel**.** The outage probability ranges from about 0.15 to 0.4. For high outage probability, the policy is similar to packet radio, with burst of high speed data when the

channel conditions are favorable. For real time applications such as voice or video that cannot tolerate the variable delays exhibited by the modulation strategy to achieve the high spectral efficiency, low outage probability is desired. However, lower outage probability requires higher power usage upper bound, which causes larger CCI. Thus, there is a trade-off between the quality of service and the frequency reuse distance.

3.4 Imperfect CSI

We now consider the case when the fed back CSI is imperfect due to delay or channel estimation error. Let α denote the actual fading channel gain and $\hat{\alpha}$ be the estimated CSI of α or related to α. The BER for imperfect CSI $\hat{\alpha}$ with transmitted power S can be obtained by evaluating the expectation of $P_e(\mathrm{M(i)}, \alpha^2 \mathrm{S/N_0})$ over α using condition *pdf* of α given $\hat{\alpha}$ as:

$$P_e^*(\mathrm{M(i)}, \mathrm{S/N_0}, \hat{\alpha}) = \int_0^\infty P_e(\mathrm{M(i)}, \mathrm{x}^2 \mathrm{S/N_0}) p_{\alpha|\hat{\alpha}}(\mathrm{x}) d\mathrm{x} \qquad (3.18)$$

Similar to the adaptive modulation for the perfect CSI (see section 3.3), the threshold for the imperfect CSI can be calculated as:

$$T_\mathrm{i} = \{\hat{\alpha} \mid P_e^*(\mathrm{M(i)}, \bar{\gamma}, \hat{\alpha}) = \mathrm{BER_{tg}}\}, \ \mathrm{i} = 1 \ldots 4 \qquad (3.19)$$

The modulation level selection for the imperfect CSI is $M = \mathrm{M(i)}$ if $T_{\mathrm{i}+1} > \hat{\alpha} > T_\mathrm{i}$. Equivalently, it can be selected as:

$$M = \max \{\mathrm{M(i)} \mid P_e^*(\mathrm{M(i)}, \bar{\gamma}, \hat{\alpha}) < \mathrm{BER_{tg}}\} \qquad (3.20)$$

The power control factor β is chosen such that $P_e^*(\mathrm{M}, \bar{\gamma}\beta, \hat{\alpha}) = \mathrm{BER_{tg}}$. The average BPS is

$$\mathrm{R_{ada}} = \sum_{\mathrm{i}=1}^{4} \log_2 \mathrm{M(i)} \int_{T_\mathrm{i}}^{T_{\mathrm{i}+1}} p_{\hat{\alpha}}(\mathrm{x}) d\mathrm{x} \qquad (3.21)$$

where $p_{\hat{\alpha}}(\mathrm{x})$ is the *pdf* of $\hat{\alpha}$. The power factor β is a function of $\hat{\alpha}$ and the average signal to noise ratio is

$$\gamma_{avg} = \int_0^{\infty} \bar{\gamma}\beta(x) p_{\hat{\alpha}}(x) dx \qquad (3.22)$$

In our analysis, we assume α is Rayleigh fading, the estimation error is modeled as white Gaussian noise and $\hat{\alpha}$ is the predicted CSI of α using linear combination of the past channel observation of α. (Channel prediction technique will be addressed in the chapter 4). Thus, α and $\hat{\alpha}$ are bivariate Rayleigh distributed and the conditional *pdf* of α given $\hat{\alpha}$ is Rician given by [24,31,59]:

$$p(\alpha|\hat{\alpha}) = p(\alpha,\hat{\alpha}) / p(\hat{\alpha}) = \frac{2\alpha}{(1-\rho)\Omega} I_0\left(\frac{2\sqrt{\rho}\alpha\hat{\alpha}}{(1-\rho)\sqrt{\Omega\hat{\Omega}}}\right) exp\left(-\frac{1}{1-\rho}\left(\frac{\alpha^2}{\Omega} + \frac{\rho\hat{\alpha}^2}{\hat{\Omega}}\right)\right) \qquad (3.23)$$

where the correlation coefficient

$$\rho = \frac{Cov(\alpha^2,\hat{\alpha}^2)}{\sqrt{Var(\alpha^2)Var(\hat{\alpha}^2)}}, \qquad (3.24)$$

$0 \le \rho \le 1$, $\Omega = E\{\alpha^2\}$, $\hat{\Omega} = E\{\hat{\alpha}^2\}$, and I_0 is the 0^{th} order modified Bessel function. From (3.20) and (3.23), the required power to employ M–QAM modulation is dependent on the parameters ρ and $\hat{\Omega}$. Nevertheless, it can be shown that when the estimated power $\hat{\Omega}$ is scaled to 1, the performance of adaptive modulation is not affected. Hence throughout the thesis, $\hat{\Omega}$ is normalized to 1 for simplicity (In practice, auto gain control (AGC) is required to track and normalized the average power of CSI), and the performance depends only on the parameter ρ. For the power control adaptive modulation method described above, we plot the BPS (3.21) vs. the correlation ρ for different SNR computed from (3.22) and for $BER_{tg} = 10^{-3}$ in Figure 3.9. The correlation $\rho = 1$ corresponds to perfect prediction, while $\rho = 0$ represents the worst case when the BPS of the adaptive modulation converges to that of the non-adaptive M-QAM for given SNR and BER_{tg}.

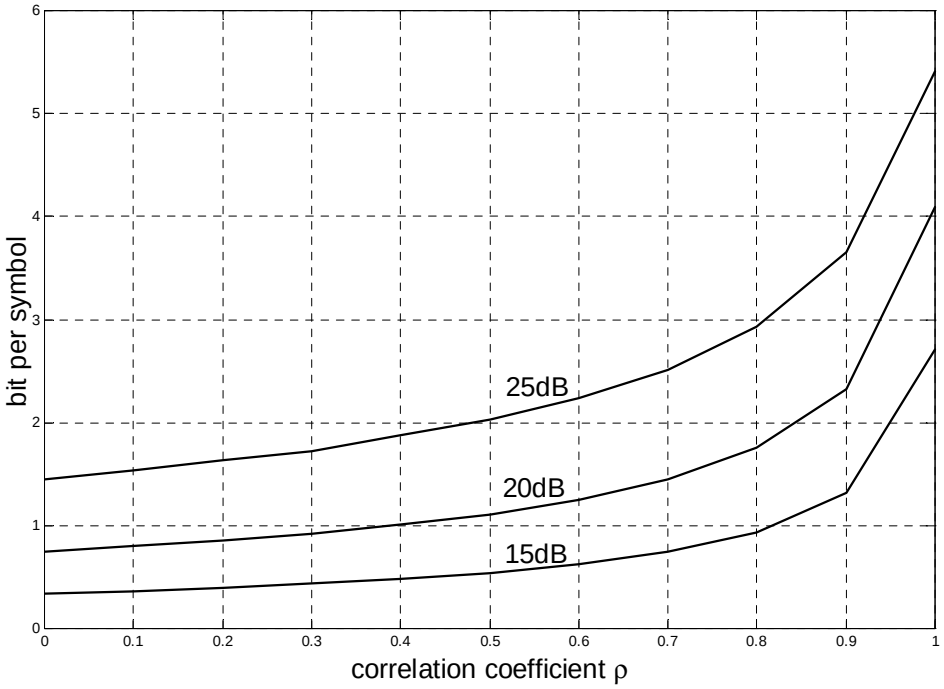


Figure 3.9 Bit per symbol vs. ρ for different SNR = S_{avg}/N_0 for power control M-QAM. Target BER=10^{-3}.

We now discuss the adaptive modulation based on the outdated CSI [24]. Assume the estimated channel gain $\hat{\alpha} = |x_1 + jy_1|$ is τ sec delay of the actual channel gain $\alpha = |x_2 + jy_2|$ due to feedback propagation delay, where $j = \sqrt{-1}$. It is shown in the Appendix C that ρ between α and $\hat{\alpha}$ can also be computed as:

$$\rho = 4\frac{u_I^2 + u_Q^2}{\Omega_1 \Omega_2} \tag{3.25}$$

where $u_I = E[x_1x_2] = E[y_1y_2]$, $u_Q = E[x_1y_2] = -E[y_1x_2]$, $\Omega_1/2 = E[x_1^2] = E[y_1^2]$ and $\Omega_2/2 = E[x_2^2] = E[y_2^2]$. For channels characterized by the correlation function $R_t(\Delta t)$ in (2.14), $u_I = R_t(\Delta t)/2$, and $u_Q = 0$. Figure 3.10 shows the correlation ρ vs. the feedback delay normalized by maximum Doppler shift f_{dm}. It is observed that even small delay causes significant performance loss. For example, for delay $f_{dm}\tau = 0.1$, the coefficient $\rho = 0.9$ and the

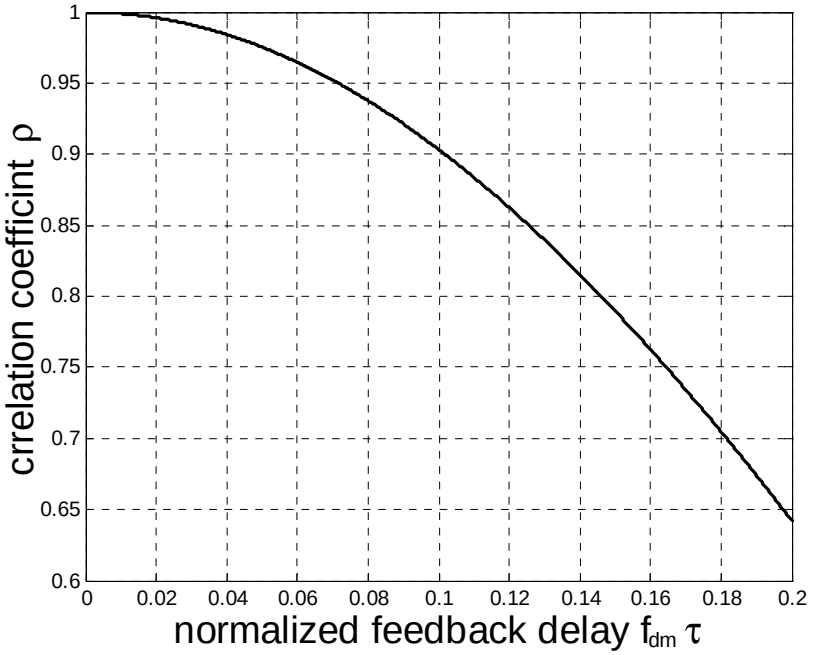


Figure 3.10 Correlation coefficient for outdated CSI. $\Omega_1 = \Omega_2 = 1$.

performance is loss is about 2 BPS relative to $\rho = 1$. Therefore, to realize the potential of the adaptive modulation method, channel prediction methods to mitigate the CSI mismatch due to feedback delay is required.

3.5 Block Loading in Time

Beside imperfect CSI, another practical consideration is the rate adaptation of the constellation size and power adaptation at the transmitter. It is easy to realize that faster AM and PC can result in larger capacity gain since the channel variations are exploited in a more accurate manner. However, hardware constraints and pulse-shaping considerations may require the modulation level to remain constant over tens or even hundred of symbols. Also, for systems that require signaling to the receiver which transmission mode is employed, the signaling overhead will significantly reduce the achievable bit rate gain. In addition, power-

amplifier linearity requirements and out-of-band emission constraints may restrict the rate at which power can be adapted [21]. Thus, block-loading techniques, where the same modulation level or power is employed for many transmission symbols, are desired.

We develop two block loading methods based on the threshold-based AM and power control described in section 3.3. Within these two methods, the power control can be categorized as per-symbol basis and frame-basis. Depending on the power control method, the largest achievable modulation level without violating the BER target and the power constraint $\bar{S}$ is used for each symbol in the frame. Let L be the frame length for block loading methods and $\{\alpha_1,\alpha_2,\ldots\alpha_L\}$ be the amplitudes of CSI sampled at the symbol rate. Assume perfect CSI and the thresholds T_i, i=1…4 are decided as in (3.11).

Per-Symbol Basis Power control (PP)

We assume the power can be continuously adjusted for each symbol in the frame. The largest achievable modulation level M is given by:

$$M = \max\{M(i) \mid \frac{1}{L}\sum_{k=0}^{L-1} \bar{S}\left(\frac{T_i}{\alpha_k}\right)^2 < \bar{S}\} \tag{3.26}$$

Once the index i is determined, the average power usage in that frame is:

$$S_f = \frac{1}{L}\sum_{k=0}^{L-1} \bar{S}\left(\frac{T_i}{\alpha_k}\right)^2 \tag{3.27}$$

Note the power control factor $\left(\frac{T_i}{\alpha_k}\right)^2$ can be larger than 1, and hence $\bar{S}$ is no longer the power usage upper-bound during the transmission.

Frame Basis Power Control (FP)

In this method, not only the modulation level but also the power remains constant for the whole frame. The modulation level is chosen as:

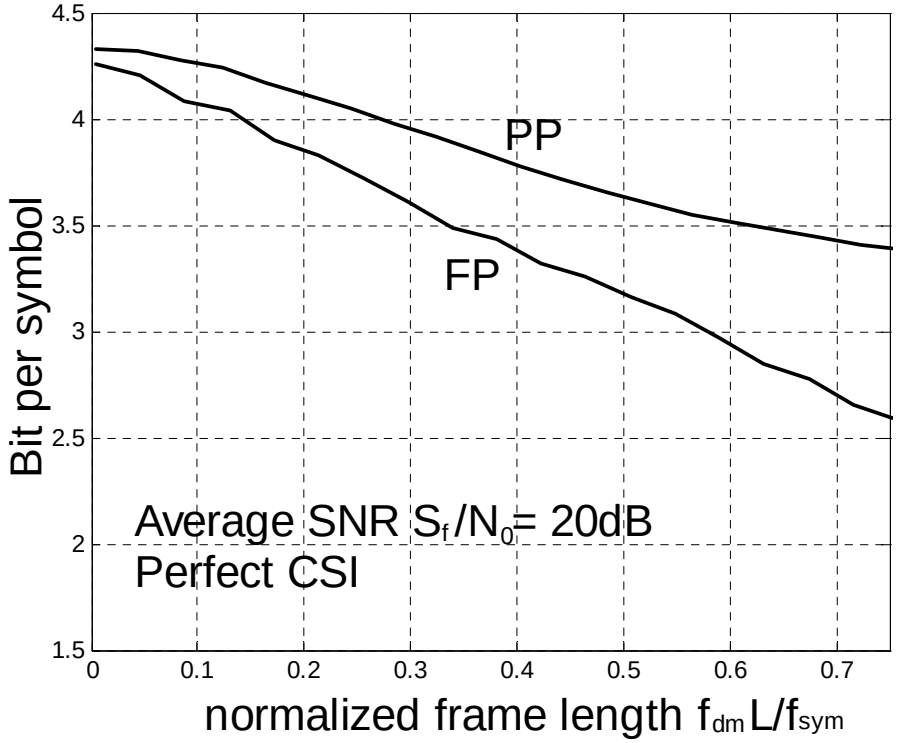


Figure 3.11 Bit per symbol for block loading adaptive modulation and power control.

$$M = \max\{\mathrm{M(i)} \mid \frac{1}{\mathrm{L}} \sum_{\mathrm{k=0}}^{\mathrm{L-1}} P_e(\mathrm{M(i)},\ \bar{\gamma}\,(\alpha_\mathrm{k})^2) < \mathrm{BER_{tg}}\} \tag{3.28}$$

The average power usage in that frame is:

$$\mathrm{S_f} = \{\bar{\mathrm{S}}\beta \mid \frac{1}{\mathrm{L}} \sum_{\mathrm{k=0}}^{\mathrm{L-1}} P_e(\mathrm{M_i},\ \bar{\gamma}\beta\ (\alpha_\mathrm{k})^2) = \mathrm{BER_{tg}}\} \tag{3.29}$$

Figure 3.11 demonstrates the numerical performance of the block loading adaptive modulation and power control. The Jakes model is used. The maximum Doppler shift f_{dm} = 100 Hz and the symbol rate f_{sym} is 25KHz. The target BER = 10^{-3}. It is observed that PP has better performance than FP since the power can be adjusted for each symbol in the frame. For frame length L = 100 ($f_{dm}L/f_{sym}$ = 0.4), the performances loss of PP and FP relative to the non-blocking loading method are about haft a bit and 0.8 bit, respectively.

CHAPTER 4

LONG RANGE CHANNEL PREDICTION AIDED BY OBSERVATIONS OF ANOTHER FADING CHANNEL

4.1 Background

As we discussed in the chapter 3, to implement adaptive transmission methods in practice, the CSI must be available at the transmitter. The CSI can be estimated at the receiver and sent to the transmitter via a feedback channel. Thus, feedback delay and overhead processing delay, and practical constraints on modulation, coding or adaptation rate have to be taken into account in the performance analysis of adaptive transmission method. For very slowly fading channels, outdated CSI is sufficient for reliable adaptive system design [24]. For fast fading that corresponds to realistic mobile speeds, however, even a small delay will cause significant degradation of performance since channel variation due to Doppler shifts usually results in a different channel at the time of transmission than at the time of channel estimation [21,24]. Even in *open-loop* [57] channels, current CSI is not sufficient since future channel conditions need to be known to adapt transmission parameters. To realize the potential of adaptive transmission methods, theses channel variations have to be predicted at least several milliseconds, or tens to hundreds of data symbols, ahead.

A novel long range channel prediction (LRP) for single carrier was proposed in [16,18,19,31]. In this chapter, we extend the long-range prediction algorithm into frequency domain. For the LRP technique, when dealing with signals that are statistically stationary, the longer the data record, the better the channel prediction that can be performed from the data. On the other hand, if the signal statistics are non-stationary, we cannot select an arbitrarily long data record to estimate the future CSI. In such a case, the length of the data record that

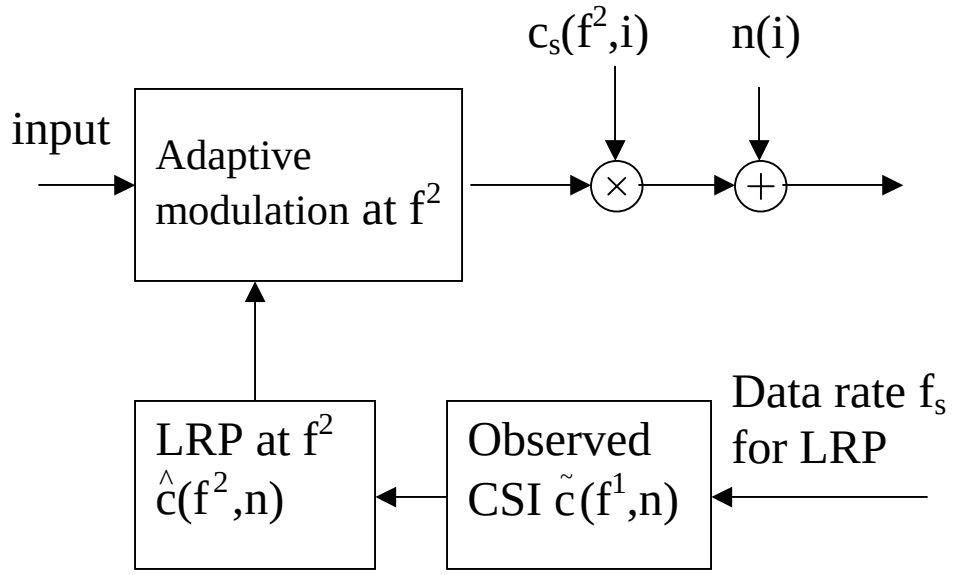


Figure 4.1 System model of adaptive modulation at carrier f^2 based on the channel observations at f^1.

we select is determined by the rapidity of the time variation in the signal statistics. Ultimately, our goal is to select as short a data record as possible that still allows us to accurately predict the future CSI. In this chapter, we first derive the MMSE LRP for the isotropic scattering channel whose correlation function is characterized by (2.12) and then develop a robust adaptive prediction method that is capable of tracking the channel statistics. Furthermore, a challenging environment is generated by the realistic physical model to test the limitation of the adaptation rate.

4.2 System Model

The discrete-time adaptive modulation system model in conjunction with the long-range prediction is illustrated in Figure 4.1. The carrier frequency of the observed CSI is f^1 and the carrier frequency of the transmitted signal is f^2. Assume narrow band signals and the

channels can be considered as *frequency non-selective*. Thus from (2.8) and (2.10), the CSI at f^1 and f^2 can be described as

$$c(f^i,t) = \sum_{n=1}^{N} A(n)\, exp\{j(2\pi fd(n)t + \phi_i(n)\} \qquad i = 1,2 \tag{4.1}$$

where for the n^{th} path, A(n) is the real amplitude and *fd*(n) is the Doppler shift. The phase difference $\phi_1(n) - \phi_2(n) = 2\pi\Delta f\tau(n)$, where $\Delta f = f^2 - f^1$ is the frequency separation, and $\tau(n)$ is the excess propagation delay distributed according to (2.13). Furthermore, the channel correlation function between f^1 and f^2 can be characterized as (2.12).

Suppose the CSI does not change within one symbol duration but varies from symbol to symbol. Let $c_s(f^2,i)$ denote the samples of the $c(f^2,t)$ with symbol interval T_{sym}, and $\tilde{c}(f^1,n)$ denote the estimate of the observed channel information at f^1 at sub-sampling rate f_s. (Note $1/(f_sT_{sym})$ is an integer.) Assume the estimation error is modeled as white Gaussian noise with power spectrum $\tilde{N}_0$. Define the observation SNE (OSNR) as

$$\text{OSNR} = \frac{\Omega_1}{\tilde{N}_0} \tag{4.2}$$

where $\Omega_1 = E[|c(f^1,t)|^2]$ is the average power of the CSI at f^1. We predict the future CSI at f^2 based on the previously observed CSI $\tilde{c}(f^1,n)$. The normalized prediction range in frequency domain is $\Delta f\sigma$. Interpolation is utilized to predict the channel coefficients $c(f^2,i)$ at the symbol rate $1/T_{sym}$.

One important feature of the LRP is its low sampling rate. That is, the sampling rate f_s is much lower than the symbol rate $1/T_{sym}$. Since the channel information at symbol rate is obtained by interpolation, from Nyquest theorem [45], the minimum data rate for LRP must be larger than twice the maximum Doppler shift to prevent aliasing. On the other hand, low

sampling rate results in longer memory span and further prediction into the future for a fixed filter length. In addition, the low sampling rate also reduces the feedback rate. However, long memory span may cause problems when severely time-variant signal have to be processed. LRP for different sampling is investigated at [19]. We use sampling rate $f_s = 5fdm$ in our numerical simulations.

4.3 MMSE Long Range Channel Prediction

The linear one-step *long range prediction* of the future channel sample $c(f^2,n)$ at frequency f^2 based on p previously observed samples at frequency f^1 is given by:

$$\hat{c}(f^2,n) = \sum_{j=1}^{p} d_j^* \tilde{c}(f^1,n-j) \tag{4.3}$$

where p is the filter length. The normalized prediction range in time is f_{dm}/f_s. The prediction error is defined by the difference

$$e(n) = c(f^2,n) - \hat{c}(f^2,n) \tag{4.4}$$

To optimize the filter design, we choose to minimize the *mean-square* value of the prediction error e(n). Define the cost function as the mean-square error (MSE):

$$J = E[|c(f^2,n) - \hat{c}(f^2,n)|^2] \tag{4.5}$$

The problem is therefore to determine the operating conditions for which *J* attains its minimum value. From [28], the necessary and sufficient condition for the cost function *J* to attain its minimum value is called *orthogonality principle*. That is, the corresponding value of the prediction error is orthogonal to each input samples that enters into the prediction of the desired response at time n, i.e.,

$$E[\tilde{c}(f^1,n-j)e^*(n)] = 0,\ j = 1,2,3,\ldots p \tag{4.6}$$

Assume the correlation function (2.12) is known. We can obtain the optimum linear filter $\underline{\mathbf{d}}_o^T = [d_{o1}, d_{o2}, \ldots d_{op}]$ (the superscript T denotes transpose) by solving a set of linear equation, called the normal equation:

$$\underline{\mathbf{d}}_o = \underline{\mathbf{R}}^{-1}\underline{\mathbf{r}} \tag{4.7}$$

where the correlation matrix of the input data:

$$\underline{\mathbf{R}} = \begin{bmatrix} R(0) & R(1) & \ldots & R(p-1) \\ R^*(1) & R(0) & \ldots & . \\ . & . & \ldots & . \\ . & . & \ldots & . \\ . & . & \ldots & . \\ R^*(p-1) & R^*(p-2) & \ldots & R(0) \end{bmatrix}, \tag{4.8}$$

$R(i) = E[\tilde{c}(f^1,n)\tilde{c}^*(f^1,n-i)] = \Omega_1 R_t(-i/f_s) + N_0\delta[i]$, where $\delta[i]$ is delta function, and $\underline{\mathbf{r}}^T = [r(1)\ r(2),\ldots..,r(p)]$ is the cross-correlation between the input data and the desired data, where $r(i) = E[\tilde{c}(f^1,n-i)c^*(f^2,n)] = R_E(i/f_s, \Delta f)$. The resulting MMSE is given by [28]:

$$\text{MMSE} = E[|e(n)|^2] = \Omega_2 - \sum_{j=1}^{p} d_{oj}^* r(j) \tag{4.9}$$

where Ω_2 is the average power of the CSI at f^2. Throughout the chapter, we normalize Ω_1 and Ω_2 to 1. This filter is well known as *Wiener filter*.

After prediction the future CSI at f^2, we perform adaptive modulation at frequency f^2 to maximize the throughput based on the predicted CSI. Assume the observed CSI $\tilde{c}(f^1,n)$ is Rayleigh distributed as in (2.5) with $\Omega = \Omega_1 + \tilde{N}_0$. Since the predicted symbol $\hat{c}(f^2,i)$ is the linear combination of the past channel observation at f^1, its amplitude is also Rayleigh distributed, and $|\hat{c}(f^2,i)|$ and $|c(f^2,i)|$ are bivariate Rayleigh distributed [59]. The correlation coefficient ρ (3.24) between $|\hat{c}(f^2,i)|^2$ and $|c(f^2,i)|^2$ is estimated and adaptive modulation can

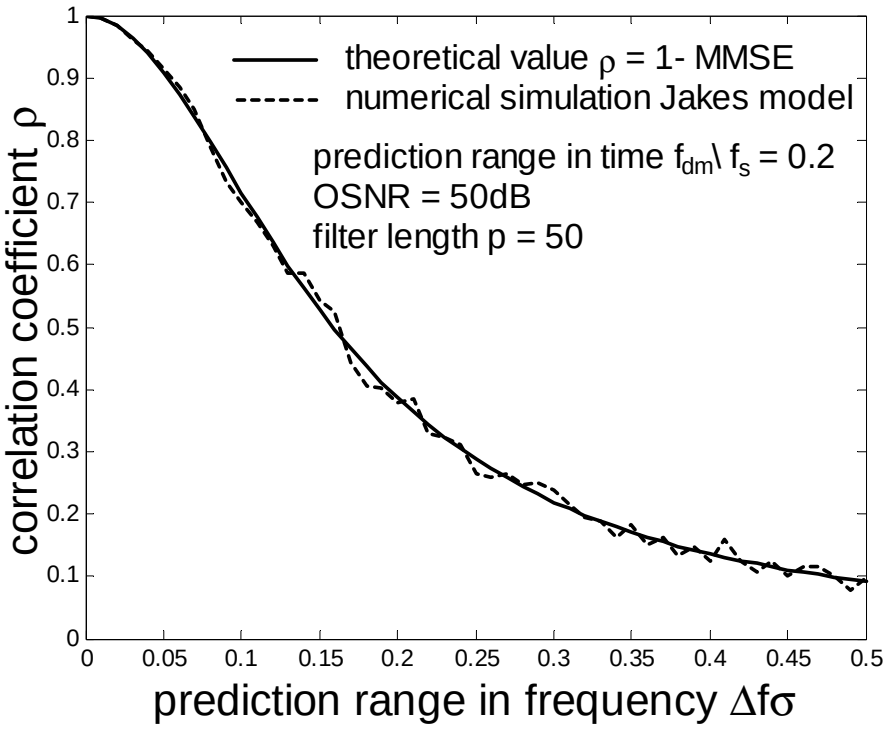


Figure 4.2 Correlation coefficient ρ vs. prediction range in frequency for theoretical MMSE and numerical simulation.

be performed as discussed in the chapter 3 (see section 3.4). In fact, when the optimal Wiener filter is employed, it is shown in the Appendix B that

$$\rho = \Omega_2 - \text{MMSE} \tag{4.10}$$

Figure 4.2 shows the theoretical correlation coefficient (4.10) vs. normalized prediction range in frequency domain filter. We also demonstrate the numerical simulation using the optimal *Wiener* filter. Jake model with N = 34 is employed to generate fading signal at f^1 and f^2, and the parameter ρ is estimated based on the predicted and actual channel state information.

4.3.1 Factors Affecting Prediction Accuracy

To achieve bit rate gain for the AM systems, accurate prediction of the CSI is required. From (4.7) and (4.9), the prediction accuracy depends on the correlation function of the input data and the correlation function between the input and the desired data. Since the

correlation function over time and frequency domain can be separated into the multiplication of two independent correlation functions (2.12), it can be readily shown that the MMSE (4.9) can be expressed as

$$\text{MMSE} = 1 - |R_f(\Delta f)|^2 \, (1 - \text{MMSE}_{\Delta f=0}) \tag{4.11}$$

where $\text{MMSE}_{\Delta f=0}$ is the MMSE with $\Delta f = 0$. The frequency domain correlation function $R_f(\Delta f)$ is dependent on the frequency separation and delay spread. Smaller frequency separation and delay spread result in larger $R_f(\Delta f)$ and thus better prediction accuracy (small MMSE). This is demonstrated in Figure 4.2. On the other hand, two factors affect the performance of single carrier channel prediction $\text{MMSE}_{\Delta f=0}$. They are noisy CSI observation and filter length.

First, the noisy channel observation degrades the prediction accuracy. When the effective SNR of the observed CSI is low, noise reduction techniques can be employed. In [19], an adaptive noise reduction method at the symbol rate is investigated, and in [27], noise reduction is implemented by suppressing out-of-band noise since the power spectrum of the fading channel is 0 outside the maximum Doppler shift (see Eq. (2.15)). In some systems, multiple steps prediction where previously predicted value is used to predict the future fading coefficients is desired for longer prediction range. In these cases, noise reduction is important since the noise will cause error propagation and make prediction accuracy unacceptable.

Second, the filter length affects the prediction performance. In general, larger filter length results in better prediction performance. As the filter length p increases, the MMSE saturation level is approached. We derive the closed form expression of MMSE for $p = \infty$ and one-step prediction (see Appendix C):

$$\text{MMSE}_{p=\infty} = R_n(0) - |R_f(\Delta f)|^2 \, [R_n(0) - r^+(0)^2 + N_0] \tag{4.12}$$

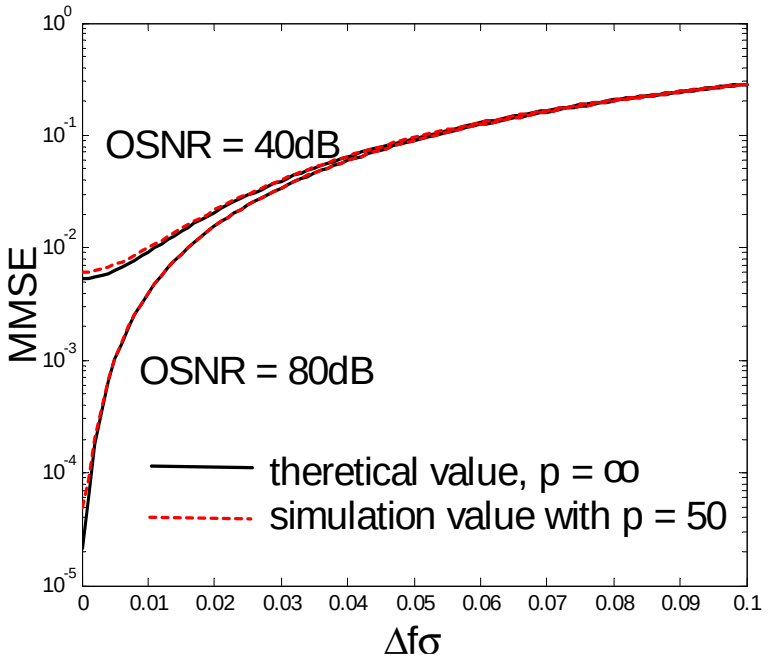


Figure 4.3 MMSE vs. normalized frequency separation $\Delta f\sigma$ for f_s =5 f_{dm}. Prediction range = 0.2/f_{dm}.

where $r^{+}(0)^2 = exp\{\frac{1}{2\pi}\int_{-\pi}^{\pi}\ln[R_w(w)+N_0]dw\}$, and $R_w(w) = \sum_{n=-\infty}^{\infty}R_n(n)exp\{-jwn\}$ is the folded power spectrum of the channel. In Figure 4.3, the theoretical MMSE of one-step prediction (4.12) is plotted vs. normalized frequency separation $\Delta f\sigma$ for different values of the signal-to-noise ratio (SNR). The sampling rate $f_s = 5fdm$ is chosen since it results in near optimal performance for LRP [19]. The prediction range is $0.2/fdm$ seconds. We also compare the MMSE of the system with filter order p = 50 (see (4.3)). We found that for p =50, the MMSE approaches the optimal case (p = ∞) for $f_s = 5fdm$. For our numerical simulation in this chapter, we employ p = 50 and the sampling rate of 500Hz assuming the maximum Doppler shift of 100Hz.

4.3.2 Backward Prediction

The form of linear prediction described in (4.3) is called *forward predictor*. We use a subset of p previously observed samples $\tilde{c}(f^1, n-1)$, $\tilde{c}(f^1, n-2)\ldots\tilde{c}(f^1, n-p)$ at f^1 to make a one step linear prediction of the sample $c(f^2, n)$ at f^2. Naturally, we may also operate on this time series in the *backward direction* [28]. In other words, we may use the subset of p samples of $\tilde{c}(f^2, n)$, $\tilde{c}(f^2, n-1)\ldots\tilde{c}(f^2, n-p+1)$ to make a prediction of the past samples $c(f^1, n-p)$, i.e.

$$\hat{c}(f^1, n-p) = \sum_{j=1}^{p} g_j^* \tilde{c}(f^2, n-j+1) \tag{4.13}$$

Similarly to (4.7), the optimal backward prediction coefficients $\mathbf{g_o} = [g_1, g_2, \ldots g_p]$ can be solved by a set of linear equation as:

$$\mathbf{R}_1 \mathbf{g_o} = \mathbf{r}_1 \tag{4.14}$$

Since we assume the fading process at f^1 and f^2 are constituted by the same Doppler shift, it can be readily shown that $\mathbf{R}_1 = \mathbf{R}$ and $r_1[i] = r^*[p-i+1]$, i=1,…,p. Note that the autocorrelation matrix **R** is *Hermitian* and *Toeplitz*. It is demonstrated in [46] that the optimal backward predictor is the reverse order complex conjugate of the optimal forward predictor, i.e.,

$$d_{oi} = g_o^*{}_{p-i+1,} \qquad i = 1\ldots p \tag{4.15}$$

We employ this relation in the following section when only finite data records are available to estimate optimal prediction filter coefficient.

4.3.3 Least Squares Estimation methods

For the LRP, we can obtain the optimum filter if the channel correlation functions are known. In practice, these channel statistics are unknown and need to be estimated from a finite data record. In this section, we use the least squares method to solve the linear prediction problem without invoking the assumption on the statistics of the input and the

desired data. The method of least squares may be viewed as an alternative to *Wiener* filter theory. Basically, *Wiener* filters are derived from ensemble average with the result that one filter is obtained for all realizations of the environment, assumed to be wide-sense stationary. On the other hand, the method of least squares is deterministic in approach. Specifically, it involves the use of time average, with the result that the filter depends on the number of samples used in the computation.

Suppose we have finite M training symbols at both carriers f^1 and f^2 during the observation interval. Define the forward prediction error at time n as:

$$f_p(f^1,f^2,n) = \tilde{c}(f^2,n) - \sum_{j=1}^{p} d_j^* c(f^1,n-j) \tag{4.16}$$

and the backward prediction error at time n as:

$$b_p(f^2,f^1,n) = \tilde{c}(f^1,n-p) - \sum_{j=1}^{p} g_j^* c(f^2,n-j+1) \tag{4.17}$$

Utilizing the fact that the optimal backward predictor is the reverse order complex conjugate of the optimal forward predictor from (4.15), the method of least squares is to minimize the cost function that consists of the sum of square errors:

$$E_{LS} = \sum_{n=p+1}^{M} |f_p(f^1,f^2,n)|^2 + \sum_{m=p+1}^{M} |b_p^*(f^2,f^1,m)|^2 \tag{4.18}$$

subject to the constraint that $d_i = g^*_{p-i+1,}$ i = 1…p. From [28], the prediction coefficient vector **d** to minimize E_{LS} can be solved be a set of linear equation:

$$\tilde{\mathbf{R}} = \mathbf{d}\tilde{\mathbf{r}} \tag{4.19}$$

where $\tilde{\mathbf{R}} = \mathbf{A}^H\mathbf{A}$, $\mathbf{A}^H = \mathbf{[FD \mid BD}^*\mathbf{]}$, and the forward data matrix **FD** and backward data matrix **BD** is:

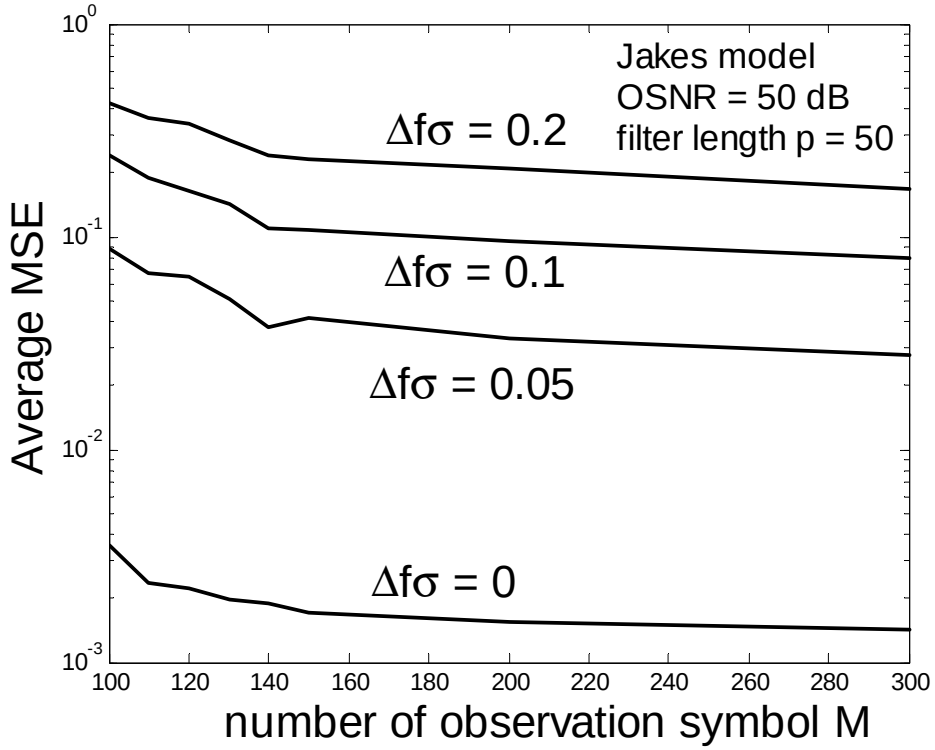


Figure 4.4 Average MSE vs. number of observation symbols for least square prediction method. normalized prediction $f_{dm}/f_s = 0.2$.

$$\mathbf{FD} = \begin{bmatrix} \tilde{c}(f^1,p) & \tilde{c}(f^1,p+1) & \cdots & \tilde{c}(f^1,K-1) \\ \tilde{c}(f^1,p-1) & \tilde{c}(f^1,p) & \cdots & . \\ . & . & \cdots & . \\ . & . & \cdots & . \\ . & . & \cdots & . \\ \tilde{c}(f^1,1) & \tilde{c}(f^1,2) & \cdots & \tilde{c}(f^1,K-p) \end{bmatrix}$$

$$\mathbf{BD} = \begin{bmatrix} \tilde{c}(f^2,K+1-p) & \tilde{c}(f^2,K-p) & \cdots & \tilde{c}(f^2,p+1) \\ . & . & \cdots & . \\ . & . & \cdots & . \\ . & . & \cdots & . \\ \tilde{c}(f^2,K-1) & \tilde{c}(f^2,K-2) & \cdots & \tilde{c}(f^1,2p-1) \\ \tilde{c}(f^2,K) & \tilde{c}(f^2,K-1) & \cdots & \tilde{c}(f^2,2p) \end{bmatrix}$$

And $\tilde{\mathbf{r}} = \mathbf{A}^H\mathbf{D}$, where $\mathbf{D}^H = [\tilde{c}(f^1,p+1), \tilde{c}(f^1,p+2) \cdots \tilde{c}(f^1,K), \tilde{c}^*(f^2,K-p), \tilde{c}^*(f^1,K-p-1) \cdots \tilde{c}^*(f^2,p)]$.

The minimum sum of squared error can be expressed as [28]:

$$E_{min,LS} = \mathbf{D}^H\mathbf{D} - \mathbf{D}^H\mathbf{A}(\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H\mathbf{D} \qquad (4.20)$$

Note that $\tilde{\mathbf{R}}$ and $\tilde{\mathbf{r}}$ in (4.19) can be considered as estimation of R and r in (4.7). In Figure 4.4, we demonstrate, for different prediction range $\Delta f\sigma$, the average prediction MSE vs. number of observation CSI samples used to obtain the least-squares based prediction coefficients. It is observed about 150 CSI samples are required to obtain near-optimal prediction performance for standard Jakes channel model.

4.4 Robust Long Range Prediction: LMS Algorithm

We derived the optimum MMSE channel prediction in (4.7) if the channel statistics, such as the time and frequency domain correlation, are known. Otherwise, the channel statistics can be estimated by least squares method in (4.19) during the observation interval. However, in the realistic mobile fading channels, the Doppler shifts in (2.8) is time-variant, and thus the model coefficients need to be updated continuously based on the observations. Since we are not able to observe the fading coefficients at frequency f^2 after observation interval, we modify our approach as follows. First, we predict future channel coefficient $c(f^1,n)$ and then use the frequency correlation function to select the transmitter parameters at f^2. The predicted CSI at f^1 are given by:

$$\hat{c}(f^1,n) = \sum_{j=1}^{p} d_j(n)^* c(f^1,n-j) \tag{4.21}$$

The coefficients $d_j(n)$ are determined using the Least Mean Square (LMS) adaptive tracking method:

$$d_j(n+1) = d_j(n) + \mu\varepsilon_n^* \hat{c}(f^1,n-j) \tag{4.22}$$

where μ is the step size and $\varepsilon(n) = c(f^1,n) - \hat{c}(f^1,n)$. This adaptive tracking can be performed since the observations at frequency f^1 are available at the transmitter [16,19]. The recursive least-squares (RLS) algorithm can also be used to improve accuracy and reduce the

observation interval [33,34]. The coefficients $\hat{c}(f^1,n)$ are interpolated to obtain predictions at the symbol rate at frequency f^1 [18].

Once $\hat{c}(f^1,n)$ is found, the adaptive modulation parameters for transmitting at f^2 at time n are selected. (Note that $\hat{c}(f^2,n)$ is not predicted directly). As explained in chapter 3 (see sec. 3.4), this procedure depends on the *pdf* of the $\alpha(f^2,n) = |c(f^2,n)|$ given $\hat{\alpha}(f^1,n) = |\hat{c}(f^1,n)|$. If we assume perfect CSI at frequency f^1 for sample n, this conditional *pdf* is determined by (3.23) with $\Omega = \hat{\Omega} = 1$ and

$$\rho = 1/(1+(2\pi\Delta f\sigma)^2) \tag{4.23}$$

In practice, this *pdf* is computed as in (3.23) using empirical estimates of $\hat{\Omega}$ and ρ and depends on the accuracy of prediction in (4.21). As discussed in the chapter 3, the performance of adaptive modulation is not affected by the value of $\hat{\Omega}$. Hence $\hat{\Omega}$ is normalized to 1 for simplicity in our numerical simulation and the performance depends only on the correlation coefficient ρ. The adaptation of ρ to the variation of the *rms* delay spread is discussed in section 4.5.

4.4.1 Performance Analysis of the LMS algorithm

The adaptive prediction algorithms involve feedback in its operation, which therefore raises the related issue of stability. A meaningful criterion is to require that

$$J(n) \rightarrow J(\infty) \tag{4.24}$$

where $J(n) = E[|\varepsilon(n)|^2]$ is the mean-square error produced by the adaptive prediction algorithm at time n, and its final value $J(\infty)$ is a constant. An algorithm that satisfies this requirement is said to be *convergent in the mean square* [28].

In general, the performance of an adaptive predictor is measured by four factors. They are convergence rate, excess mean square error, tracking ability, and complexity. First,

the convergence rate tells how fast the adaptive algorithm can reach its steady state. Second, the excess mean square error $J_{ex}(\infty)$ is the difference between the final value $J(\infty)$ and the minimum value J_{min} attained by the Wiener solution (4.9). This difference represents the price paid for using the adaptive stochastic mechanism to control the filter coefficients instead of the expected value of the channel statistics. However, it is important to realize that this value is under designer's control. Third, tracking ability, contrasted with convergence, is a steady state phenomenon. For an adaptive predictor to examine its tracking capability, it must first pass from the transient mode to the steady state mode. We use the physical channel model to generate non-stationary environment to test the tracking ability of our adaptive prediction algorithm. Note that convergence rate and tracking ability are two different properties. In particular, an adaptive prediction algorithm with good convergence properties does not necessarily possess a fast tracking capability, and vice versa [28]. Finally, complexity measures the cost of the system.

The transient behavior of the mean square error $J(n)$ of the LMS algorithm and the excess mean square error $J_{ex}(\infty)$ are determined by the step size μ and the correlation matrix $\underline{\mathbf{R}}$ (4.7) of the input data. By unitary similarity transformation [28], the correlation matrix can be decomposed as:

$$\mathbf{Q}^H \underline{\mathbf{R}} \mathbf{Q} = \mathbf{\Lambda} \tag{4.25}$$

where $\mathbf{\Lambda}$ is a diagonal matrix consisting of the eigenvalue λ_i of the correlation matrix $\underline{\mathbf{R}}$, and $\mathbf{Q}$ is the unitary matrix consisting of the eigenvectors associated with these eigenvalues. The transient behavior $J(n)$ is derived in [28] as:

$$J(n) = J_{min} + \sum_{i=1}^{p} \lambda_i x_i(n) \tag{4.26}$$

where $x_i(n)$ can be determined iteratively as:

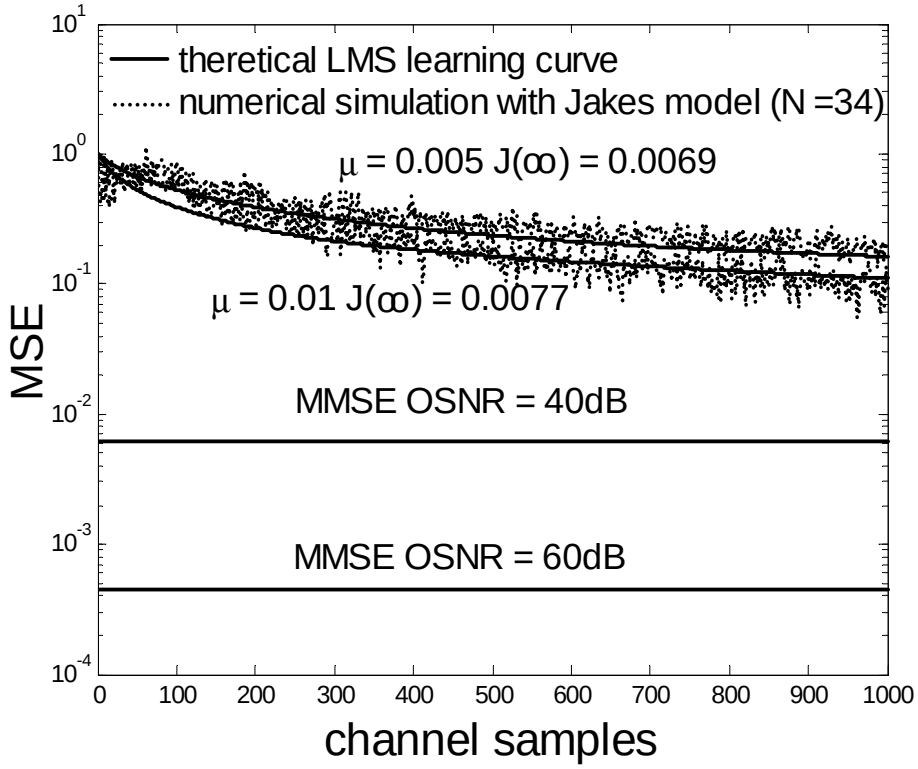


Figure 4.5 LMS algorithm learning curve. Normalized Prediction range f_{dm}/f_s = 0.2. filter length p =50.

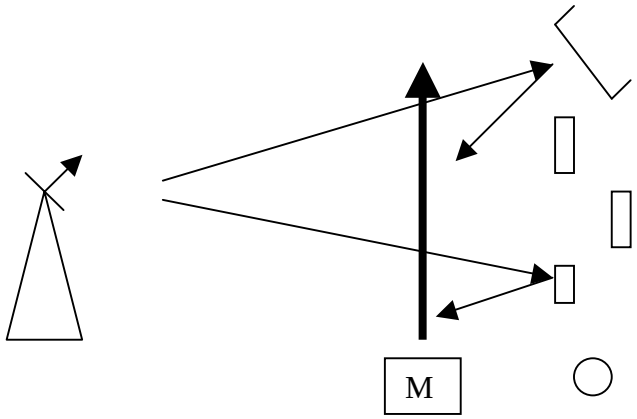
$$x_i(n+1) = (1 - \mu\lambda_i)^2 x_i(n) + \mu^2 J_{\min}\lambda_i \tag{4.27}$$

The initial value of $x_i(0)$ is the i^{th} diagonal component of the matrix $\mathbf{Q}^H\mathbf{d}_0\mathbf{d}_0{}^H\mathbf{Q}$. The selection of step size μ is a trade−off between the convergence rate and the excess MSE. We obtain good results with μ from 0.005 to 0.01. Figure 4.5 shows the theoretical LMS learning curve in (4.26) and the theoretical MMSE in (4.9). It is observed that the LMS algorithm converges slowly with significant excess mean square error. We also found that while OSNR determines the performance of the theoretical MMSE, it does not significantly effect the learning curve of the LMS algorithm on the fading channel characterized by the correlation function $R_t(\Delta t)$ in (2.14). We also demonstrate the numerical simulation using Jakes model for comparison. In spite of its slow convergence rate, LMS is a very simple algorithm to implement. It does not require measurements of the pertinent correlation functions, nor does

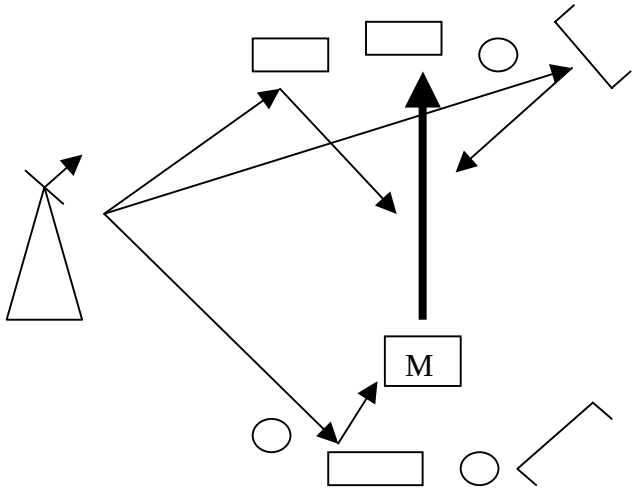
it require matrix inversion. Indeed, it is the simplicity of the LMS algorithm that has made it the standard against which other adaptive filtering algorithms are benchmarked. The LMS algorithm requires 2p+1 complex multiplication and 2p complex additions per iteration. In other words, the computational complexity of the LMS algorithm is O(p) [28].

4.5 Channel Generation for Realistic Physical Model

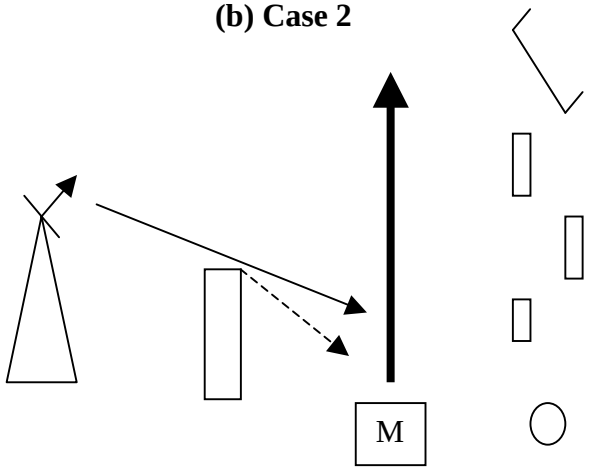
The prediction accuracy of our system depends not only in normalized time difference but also in the normalized frequency separation $\Delta f\sigma$. For adaptive modulation algorithm, we assume the *rms* delay spread σ is fixed while in realistic fading channels, it is time variant. Thus, σ needs to be estimated and updated at the transmitter. We use the physical model to generate different time-variant *rms* delay spread environments to investigate the limits on the speed of the adaptation. The geometries for generating varying delay spreads are shown in Figure 4.6 (a),(b), and (c), and the delay spread for each case is shown in Figure 4.7. The distance between the transmitter and the mobile is 1Km, the sampling rate is 500 Hz, the speed of the mobile is 30m/s and the carrier frequency is 900MHz. This corresponds to the maximum Doppler shift 100 Hz. In the case 1, the direction of the mobile is approximately parallel to the scattering objects. In the case 2, the scattering objects are arranged to be perpendicular to the moving direction. Figure 4.7 shows that the delay spread for the case 1 remains approximately constant since the propagation paths do not change significantly. On the other hand, the delay spread for the case 2 changes more rapidly relative to the case 1 due to the significant variation of the delay spread. This is consistent with the results in [29] that the CSI is slowly variant and easy to track when the scatterers are parallel to the direction of the mobile; and the CSI changes rapidly and is hard to track if the mobile moves toward the scattering objects.



(a) Case 1



(b) Case 2



(c) Case 3

Figure 4.6 Geometries of three physical modeling cases.

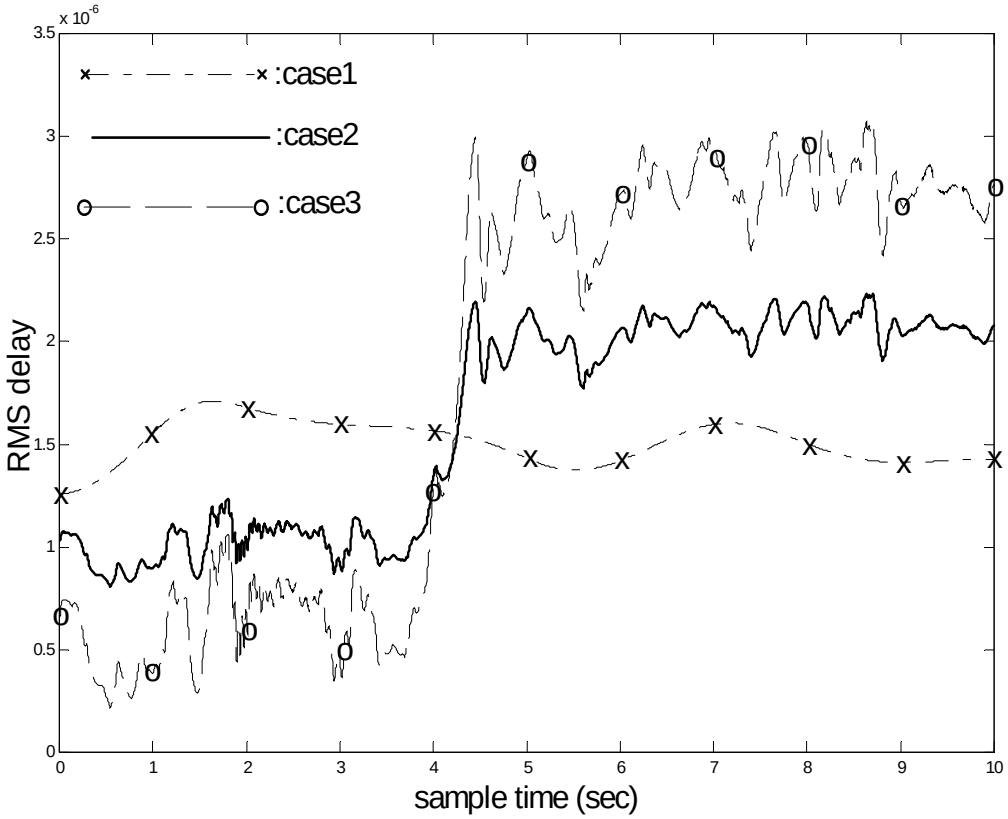


Figure 4.7 The variation of the *rms* delay spread σ for typical (case 1) and challenging cases (cases 2 and 3).

In the last case, the shadow fading is modeled. This situation can occur in an urban area where there exist large buildings. In Figure 4.7 (c), the mobile is shadowed by buildings in the beginning and can only receive signals by diffraction phenomenon (dashed line). When the mobile moves ahead, it can receive the signals directly from the source (LOS), resulting a significant change of the impulse response and the rms delay spread. Due to the severe variation of the delay spread, the frequency domain correlation also changes significantly. We call the third case the challenging case since accurate estimate and rapid track of the rms delay spread is required to maintain the performance of our adaptive modulation system.

Due to the capability of the physical model to capture the properties of a real mobile channel, the obtained datasets from the physical model usually exhibit, in addition to small scale fading (multipath fading), large scale fading (shadowing, power loss over distance). To compare the system performance over the physical model with that over the stationary Jakes model, it is important to separate the datasets from large scale fading. It is discussed in [6] that the small-scale fading can be approximately obtained by calculating $A(x)/\bar{A}(x)$, where $A(x)$ is the envelope of the obtained data, and $\bar{A}(x)$ a small-area-average calculated by:

$$\bar{A}(x) = \frac{1}{2W} \int_{-W}^{W} A(x + t)\, dt \tag{4.28}$$

where W is the window length used to calculate the average at time x. The small-area-average should be calculated on the order of 10m [6]. For example, W = 0.3 sec can be chosen if the speed of the vehicle is 30m/s. In a real mobile channel, this technique can be used in the auto-gain-control (AGC) to track the mean value of the signal power.

4.6 Numerical Simulation

We use the Jakes model and the aforementioned physical model to validate the performance of the continuous power discrete rate adaptive M–QAM discussed in the chapter 3 aided by the LRP. The bit rate and average SNR is calculated as in (3.21) and (3.22). The maximum Doppler shift of 100 Hz is used in both models. The target BER = 10^{-3}. The fading signal is sampled at the rate of 500Hz for the LRP. The observation interval is 150 samples, the OSNR in the observations is 80dB, the symbol rate is 25ksymbol/s, and the modulation-switching rate is the same as the symbol rate. Interpolation is utilized to predict the channel coefficients at the symbol rate. The prediction range is 2ms.

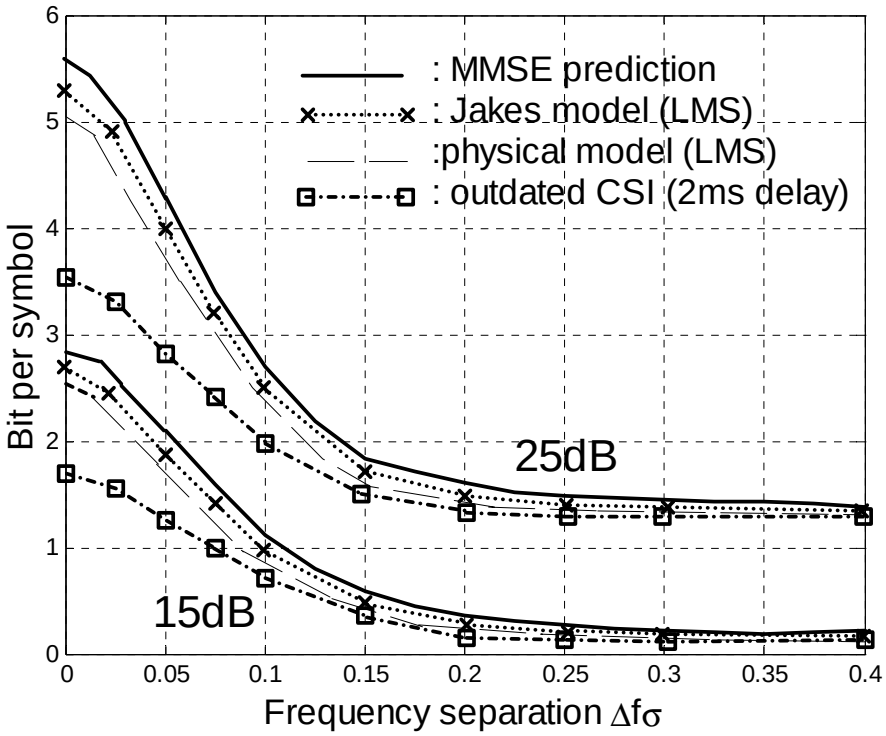


Figure 4.8 BPS vs. normalized frequency separation for different prediction techniques. f_{dm} = 100Hz. Prediction range is 2ms.

In Figure 4.8, we plot BPS vs. normalized frequency separation $\Delta f\sigma$ for the ideal (non-adaptive) MMSE filter (4.7) and the robust method using the LMS algorithm with the step size 0.005. The parameters ρ in (3.23) are estimated during the observation interval for both data sets and are used throughout the transmission. For example, for $\Delta f = 0$, the estimated $\rho = 0.983$ for the Jakes model and $\rho = 0.965$ for the physical model (case 1). The bit rate loss is less than half a bit for non-stationary data generated by the physical model relative to the stationary case. We also investigated the BPS under the assumption that prediction is perfect at frequency f^1. We observe that the performance of the robust algorithm is very close to this ideal case as well as to the performance of the ideal MMSE algorithm. This is consistent with the results in [31] where it is shown that when observations and the predicted samples are at the same frequency, performance of adaptive modulation

aided by robust (adaptive) LRP closely approximates the ideal performance with perfect CSI. Hence the robust method is near-optimal and has the ability to adapt transmission parameters to the time-variant channel conditions. Moreover, for given σ, the theoretical value of the parameter ρ in (4.23) can be utilized in the selection of thresholds when robust prediction is used.

The performance of the adaptive modulation using the outdated CSI for the Rayleigh fading channel with the correlation function (2.12) is also shown in Figure 4.8. To alleviate the mismatch of the delayed and future CSI, a novel approach to calculate thresholds based on the delayed CSI was studied in [24]. A similar method is employed here. A single observation $c(f^1,n-1)$ is used instead of the estimate $\hat{c}(f^1,n)$ in (4.21) to compute the modulation parameters at frequency f^2. We found that even very small delay causes significant loss of the bit rate for fast vehicle speeds when accurate long range prediction is not utilized. For example, for $\Delta f\sigma=0$ and τ = 2ms, the loss is 1 to 2 BPS. Thus, accurate LRP is required to achieve the bit rate gain of adaptive modulation for fast vehicle speeds and realistic delays.

Figure 4.8 also shows that adaptive modulation is primarily beneficial when the normalized frequency separation $\Delta f\sigma$ does not significantly exceed 0.1. For example, for $\Delta f\sigma = 0.1$, about 17dB is required to obtain 1 BPS for adaptive M-QAM as opposed to 24dB for non-adaptive signaling (BPSK). As $\Delta f\sigma$ approaches 0.4, the bit rate of adaptive modulation approximates that of non-adaptive transmission. Hence the frequency separation and the multipath delay (or the coherence bandwidth) are the factors that determine the performance of the proposed adaptive modulation method. The typical values of σ are on the order of microseconds in outdoor mobile radio channel [51]. Suppose $\Delta f\sigma = 0.1$ and $\sigma = 1\mu$

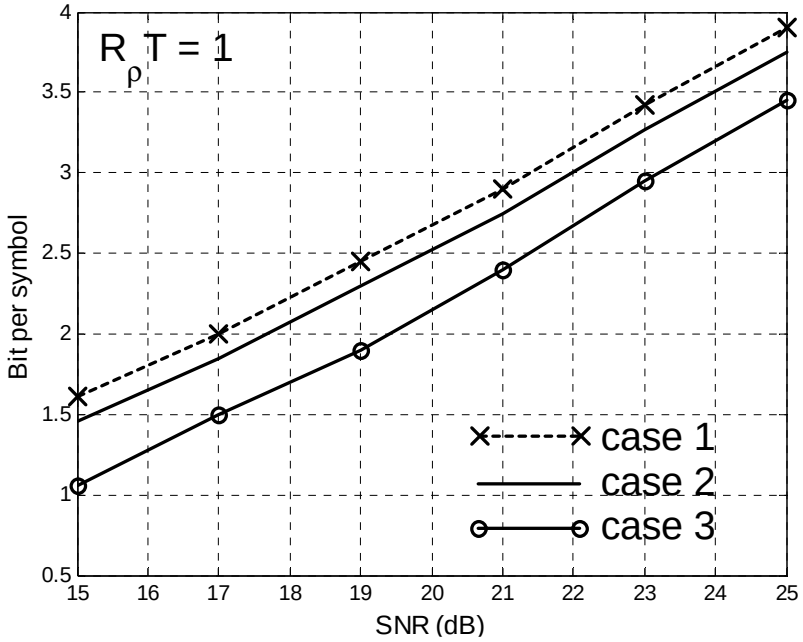


Figure 4.9 BPS vs. SNR for $R_\rho T$ =1. Δf = 50KHz.

sec. Then the frequency separation $\Delta f = 100$KHz. This means that two channels can be separated by 100KHz and still benefit from the proposed adaptive transmission method.

Another practical consideration is the adaptation of the parameter ρ at the transmitter as a function of the variation of the *rms* delay spread σ. To investigate the limits on the speed of adaptation, we use the physical model to generate challenging and typical scenarios shown in Figure 4.9. We investigate the performance of adaptive modulation on these channels during the T=1 sec interval when σ varies rapidly in cases 2 and 3 (from 3.5 to 4.5 sec in Figure 4.7). The variation of the *rms* delay spread is approximately from 0.7 to 2.3 μs and 0.4 to 2.6 μs for cases 2 and 3, respectively. The target BER is 10^{-3} and the power is adjusted to maintain the target BER to compensate for the mismatch of the *rms* delay. The parameter ρ is updated at the rate R_ρ Hz. Figure 4.9 shows the BPS vs. SNR for the normalized adaptation rate $R_\rho T = 1$, i.e. the value of ρ is not updated during the interval T. There is about 2dB loss

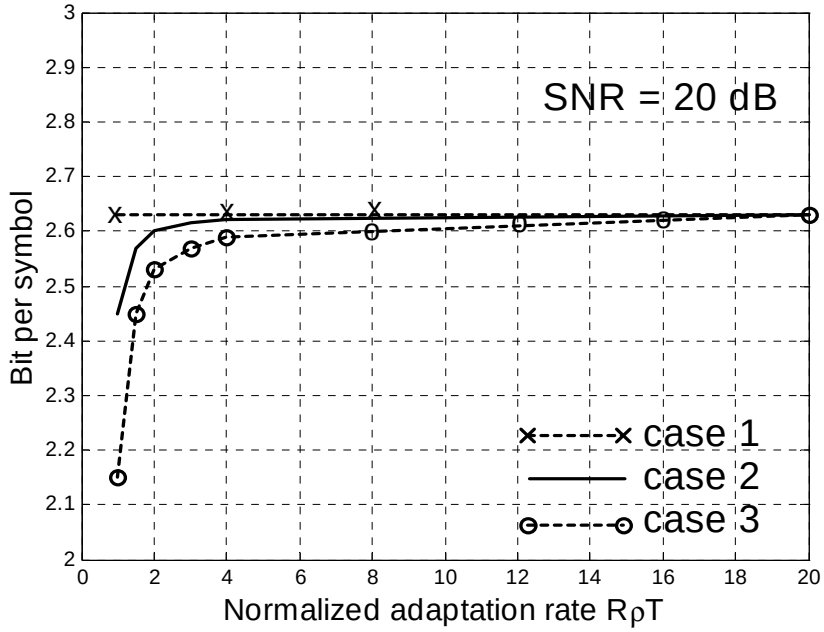


Figure 4.10 BPS vs. normalized adaptation rate $R_\rho T$. SNR = 20dB. $BER_{tg}=10^{-3}$. Δf = 50KHz.

for the challenging case 3 relative to the *rms*-invariant case 1. To improve performance, the correlation ρ needs to be tracked and updated more frequently. Figure 4.10 illustrates that there is significant performance loss for the challenging case if $R_\rho T < 2$. By analyzing datasets produced by the physical model, we concluded that the variation of the *rms* delay (and ρ) is typically slow and tracking of the correlation ρ does not result in significant additional computational and feedback load. The required rate of update of the parameter ρ is significantly slower than the low sampling rate for predicting at frequency f^1 in (4.21). Thus, the proposed robust prediction method based on the observations at a different carrier is feasible, but infrequent update of the time-variant frequency correlation is required to satisfy the adaptive transmission performance criterion (e.g. the BER_{tg}).

In this paper, the assumption of the exponentially distributed propagation delay (2.13) results in the relationship of the parameter ρ and the *rms* delay spread σ that is approximated

by (4.23). If the distribution of the propagation delay is different, this relationship will change. For example, for the uniform distribution, the coherence bandwidth and ρ are reduced for a given σ, and hence the performance of the prediction in the frequency domain and the bit rate are degraded relative to the exponentially distributed excess delay. Since we directly estimate the correlation ρ from the dataset, our algorithm is robust to the variation in the distribution of the excess delay.

CHAPTER 5

ADAPTIVE OFDM SYSTEM AIDED BY LONG RANGE PREDICTION

5.1 Background

High data rate communications are limited not only by noise, but often more significantly by the inter symbol interference (ISI) [45] due to the time dispersive wireless communication channel. As a general rule, the effects of ISI on the transmission error statistics are negligible if the delay spread is significantly shorter than the duration of the transmitted symbol. For higher symbol rate transmissions, mechanisms such as channel equalizers [44] must be implemented in order to combat the effect of ISI. Significant research efforts were invested to develop such channel equalizer, and most wireless systems in operation use equalizer to combat ISI.

There is, however, an alternative approach called Orthogonal Frequency Division Multiplexing (OFDM) for transmitting data over a multipath channel. The basic principle of OFDM is to split a high-rate data stream into a number of lower rate streams that are transmitted simultaneously over a number of subcarriers. The relative amount of dispersion in time caused by the multipath delay spread is decreased due to the increased symbol duration for the lower rate parallel subcarriers, resulting in relatively simple receiver structure compared with single carrier transmission in frequency selective fading channels. In addition, OFDM can significantly enhance the capacity by adapting the data rate per subcarrier according to the signal-to-noise ratio of that particular subcarrier.

In this chapter, we first review the modulation and demodulation process of the OFDM signal. Next, we develop the long-range prediction algorithms for OFDM systems.

An ideal MMSE method that utilizes previous observations in time and frequency domain, and robust adaptive LRP algorithms are developed and compared. The LRP is utilized in adaptive bit and power allocation for the OFDM system. Statistical model of the prediction error is created and used in the design of reliable adaptive modulation. In addition, several methods that significantly reduce the feedback load for mobile radio AOFDM systems are developed and compared.

5.2 Modulation and Demodulation for OFDM Signals

In a classic parallel data system, the total signal frequency band is divided into K non-overlapping frequency subcarriers. Each subcarrier is modulated with a separate symbol and then the K subcarriers are frequency-multiplexed. While it is desirable to avoid spectral overlap of channels to eliminate interchannel interference (ICI), this might lead to inefficient use of the available spectrum. To cope with the inefficiency, the ideas proposed from the mid-1960s [8,52] were to use parallel data and FDM with overlapping subcarriers.

As an example, Figure 5.1 shows four subcarriers $S_1(t)$ to $S_4(t)$ from one OFDM signal with different phases and amplitudes in a symbol interval T_m ($t = 0$ to T_m). To maintain orthogonality between these subcarriers from 0 to T_m ($\int_0^{T_m} S_i(t)S_j(t)dt = 0$ if $S_i(t)$ and $S_j(t)$ are orthogonal for $i \neq j$), each subcarrier has exactly an integer number of cycles in the interval $[0,T_m]$, and the number of cycles between adjacent subcarriers differs by exactly one. An additional guard time (GT) from $-T_g$ to 0 is inserted to prevent multipath induced interchannel interference at the receiver. For instance in the Figure 5.1, $S'_1(t)$, the delayed version of $S_1(t)$, maintains the orthogonality to the other subcarriers $S_2(t)$, $S_3(t)$, and $S_4(t)$ without causing any interference and vice versa from $t = 0$ to T_m due to the guard time. On

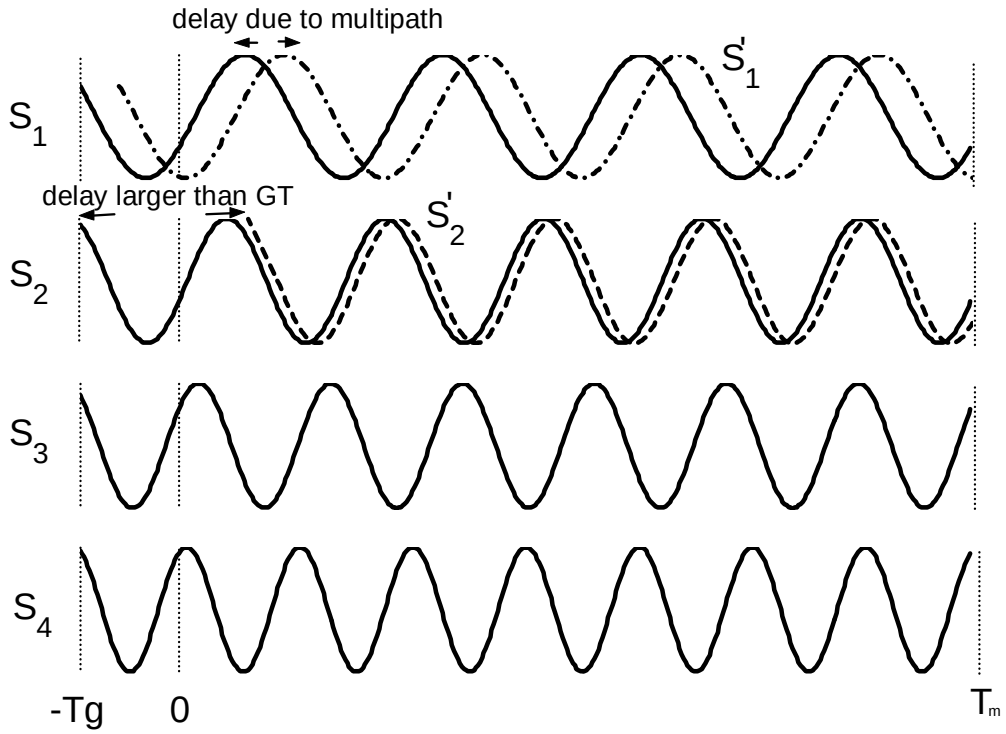


Figure 5.1 Example of four subcarriers within one OFDM symbol

the other hand, $S'_2(t)$, the delayed version of $S_2(t)$, causes interference to other subcarriers since the excess propagation delay is larger than the GT. Thus, the GT is chosen larger than the expected maximum excess propagation delay, such that the multipath components from one symbol cannot interfere with other symbols.

From the above discussion, an OFDM signal with K subcarriers, symbol interval T_m and guard time T_g at carrier frequency f_c can be represented as:

$$s(t) = \sum_{i=0}^{K-1} A_i \cos(2\pi(f_c + \frac{i}{T_m})t + \varphi_i), \; -T_g \le t \le T_m \qquad (5.1)$$

where for the i^{th} subcarrier, A_i is the (real) amplitude, φ_i is the phase. The equivalent complex baseband notation can be expressed as:

$$s_l(t) = \sum_{i=0}^{K-1} s_i \exp\{j2\pi\frac{i}{T_m}t\}, \; -T_g \le t \le T_m \qquad (5.2)$$

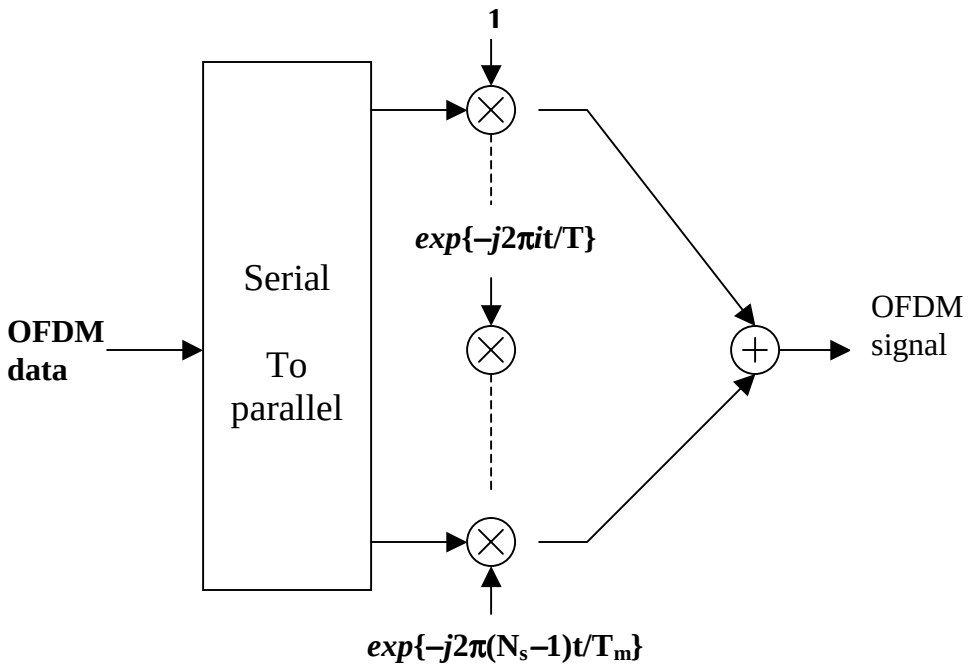


Figure 5.2 OFDM modulator

where $s_i = A_i exp\{j\varphi_i\}$ are complex transmitted symbols. Figure 5.2 shows the operation of the OFDM modulator in a block diagram. Unfortunately, for a large number of channels, the array of sinusoidal generators and coherent demodulators required in parallel systems become unreasonablely expensive and complex. However, it is shown in [63] that (5.2) can be approximately generated by sampled data from $s_l(t)$ with sampling interval T_m/K and a lowpass filter. The sampled data s[*n*] can be expressed as:

$$s[n] = \sum_{i=0}^{K-1} s_i \, exp\{j2\pi \frac{in}{K}\}, \; -N_g \le n < K \tag{5.3}$$

where N_g is the number of samples in the GT. Notice that the sampled data {s(*n*), *n* = 0,…,K–1} is nothing more than the inverse discrete Fourier transform (IDFT) of K input complex data {s_i, *i* = 0…*Ns* –1}. And the data set {s(*n*), *n* = $–N_{g,\ldots,}–1$} are exactly the

duplicate of the data set{s(n), n = K–N_g,…,K–1}. Therefore, they are called the *cyclic prefix* (CP) code.

Suppose the channel is time-invariant and let $h_l(\tau)$ denote the equivalent complex lowpass channel impulse response. Assume the GT is larger than the maximum delay spread, i.e., $h_l(\tau) = 0$ for $\tau > T_g$. The received OFDM signal $r_l(t)$ can be expressed as:

$$r_l(t) = \int_0^{\infty} h_l(\tau)\, s_l(t-\tau)d\tau = \sum_{i=0}^{K-1} \int_0^{\infty} h_l(\tau) exp\{-j2\pi \frac{i}{T_m}\} d\tau\, s_i\, exp\{j2\pi \frac{i}{T} t\},\ -0 \le t \le T_m \quad (5.4)$$

To demodulate the signal s_i, $r_l(t)$ is sampled at sampling rate T_m/K, and the sampled data $\{r_l[n], n = 0,\ldots,K-1\}$ are taken Discrete Fourier Transform (DFT). The resulting symbol at i^{th} subcarrier can be given as

$$u[i] = s_i \int_0^{\infty} h_l(\tau) exp\{-j2\pi \frac{i}{T_m}\} d\tau = c[i]\, s_i,\ i = 0,\ldots,K-1 \quad (5.5)$$

It is observed that the desired signal s_i is effected by a complex channel gain c(i), which is the frequency response at subcarrier i. Thus, the signal can be demodulated by a simple one-tap complex equalizer.

5.2.1 OFDM over Mobile Fading Channel

As we discussed in the previous section, OFDM signal can be demodulated without any ICI as long as the guard time is larger than the maximum excess delay spread and the channel impulse response is time-invariant. However, the impulse response is time-variant for mobile radio channels and a more interesting issue is the effect of Doppler shift on the OFDM receiver performance. Assume the GT is always longer than the maximum delay spread, so that the effects of ISI are totally removed and the blocks can be analyzed

independently. From (2.8) and (2.10), the equivalent lowpass complex fading coefficients at K subcarriers $f^1,\ldots,f^K$, $f^1<f^2<\ldots<f^k$, where $|f^i-f^j|<<$ the carrier frequency f_c, can be closely approximated as [35]:

$$c(f^i,t)=\sum_{n=1}^{N} A(n)exp\{j(2\pi f_d(n)t+\phi_i(n)\}, \quad i = 1,2\ldots K \tag{5.6}$$

Ignoring AWGN, the equivalent complex lowpass received OFDM signal, after multipath fading channel, can be characterized as:

$$r_l(t) = \sum_{m=1}^{N}\sum_{i=0}^{K-1} A_m(i)\, s_i\, exp\{j2\pi(\frac{i}{T_m} + f_d(m))t\},\ 0 \le t \le T_m \tag{5.7}$$

where N is the total number of the path, and for the m^{th} path, $A_m(i) = A(m)exp\{j\phi_i(m)\}$is the complex fading gain, and $f_d(m)$ is the Doppler shift. To demodulate at the receiver, $r_l(t)$ is sampled at the sampling interval T_m/K, resulting in:

$$r_l[n] =\sum_{m=1}^{N}\sum_{i=0}^{K-1} A_m(i)\, s_i\, exp\{j2\pi(\frac{i}{T_m} + f_d(m))n\frac{T_m}{N_s}\},\ n = 0,\ldots,K-1$$

$$= \sum_{m=1}^{N}\sum_{i=0}^{K-1} A_m(i)\, exp\{j2\pi f_m\frac{nT_m}{K}\}\ s_i\, exp\{j2\pi\frac{inT_m}{K}\},\ n = 0,\ldots,K-1 \tag{5.8}$$

Let $c[k, n] = \sum_{m=1}^{N} A_m(k)\ exp\{j2\pi f_m\frac{nT_m}{K}\}$, which is the complex fading channel gain at k^{th} subcarrier with sampling interval T_m/K. Define $X_i(n) = c[i, n]\ s_i\ exp\{j2\pi\frac{inT_m}{K}\}$, then (5.8) can be rewritten as:

$$r_l(n) = \sum_{i=0}^{K-1} X_i(n) \tag{5.9}$$

Let $W = exp\{-j\frac{2\pi}{K}\}$ and $\{\tilde{s}(n), n = 0\ldots K-1\}$ denote the symbols after taking DFT of $\{r_l(n), n = 0\ldots K-1\}$. Then the i^{th} component, $\tilde{s}(i)$, can be expressed as:

$$\tilde{s}(i) = \frac{1}{K}[(W^i)^0\ (W^i)^1 \ldots\ (W^i)^{K-1}] \begin{bmatrix} r_l[0] \\ r_l[1] \\ \cdot \\ \cdot \\ r_l[K-1] \end{bmatrix} = \frac{1}{K}\ [(W^i)^0\ (W^i)^1 \ldots\ (W^i)^{K-1}] \sum_{i=0}^{K-1} \begin{bmatrix} X_k[0] \\ X_k[1] \\ \cdot \\ \cdot \\ X_k[K-1] \end{bmatrix}$$

$$= \frac{1}{K}\ [(W^i)^0\ (W^i)^1 \ldots\ (W^i)^{K-1}] \begin{bmatrix} X_i[0] \\ X_i[1] \\ \cdot \\ \cdot \\ X_i[K-1] \end{bmatrix} + \frac{1}{K} \sum_{\substack{k=0 \\ k \neq i}}^{K-1} [(W^i)^0\ (W^i)^1 \ldots\ (W^i)^{K-1}] \begin{bmatrix} X_k[0] \\ X_k[1] \\ \cdot \\ \cdot \\ X_k[K-1] \end{bmatrix}$$

$$= \frac{1}{K} \sum_{m=0}^{K-1} c[i, m]\ s_i + \frac{1}{K} \sum_{\substack{k=0 \\ k \neq i}}^{K-1} s_k\, c_{k,i} \qquad (5.10)$$

where $c_{k,i} = \sum_{m=0}^{K-1} (W^i)^m c[k, m]$ is the ICI from k^{th} subcarrier to the i^{th} subcarrier. By comparing (5.10) with (5.5), it is observed that OFDM signals over multipath fading channel not only result in different CSI $\frac{1}{K} \sum_{m=0}^{K-1} c[i, m]$ but introduce a ICI. Assume the channel gain $c[i, m]$ does not change within one OFDM symbol interval T_m, i.e., $c[i, m]$ is a constant for $m = 0\ldots K-1$, thus the channel gain $\frac{1}{K} \sum_{m=0}^{K-1} c[i, m]$ at subcarrier i can be approximately modeled as the samples of the time-varying frequency selective channel in (5.6) with the time domain interval T_m. The effect of the ICI is not considered in this thesis, although in practice it can be significant and has to be addressed in the overall system optimization. [47].

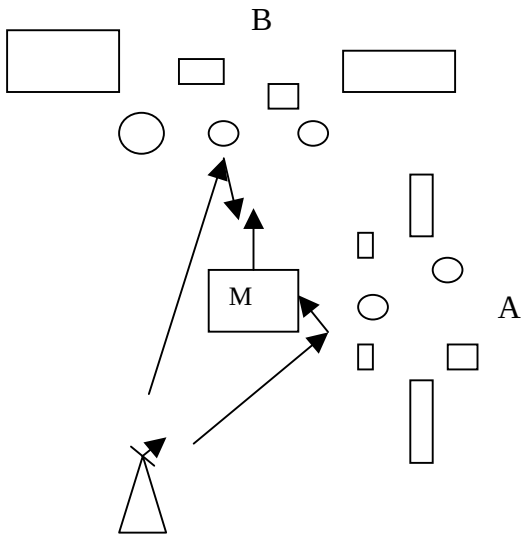


Figure 5.3 Geometries of the physical modeling.

5.2.2 Physical Model

We employ the physical model to generate the realistic multicarrier channel. The geometry for generating the model data set used in this chapter is shown in Figure 5.3. The scattering objects are arranged approximately parallel (group A) and perpendicular (group B) to the direction of the mobile. The fading amplitude is shown in Figure 5.4. The maximum Doppler shift is 100Hz, and the sampling rate is 500Hz. For the first 500 samples, the CSI is influenced significantly by the multipath from group A, whereas the CSI is dominated by the multipath from group B for the following 500 samples due to the reversal of signal amplitudes from the relative distance to the scattering objects. Taking into account the direction of the mobile, we see that the Doppler shift rapidly changes in both magnitude and sign, and the dominant amplitudes and phases undergo significant variation during the transition period from samples 500 to 700. We use this transition interval to test the robustness of the LRP to parameter variation in section 5.3.

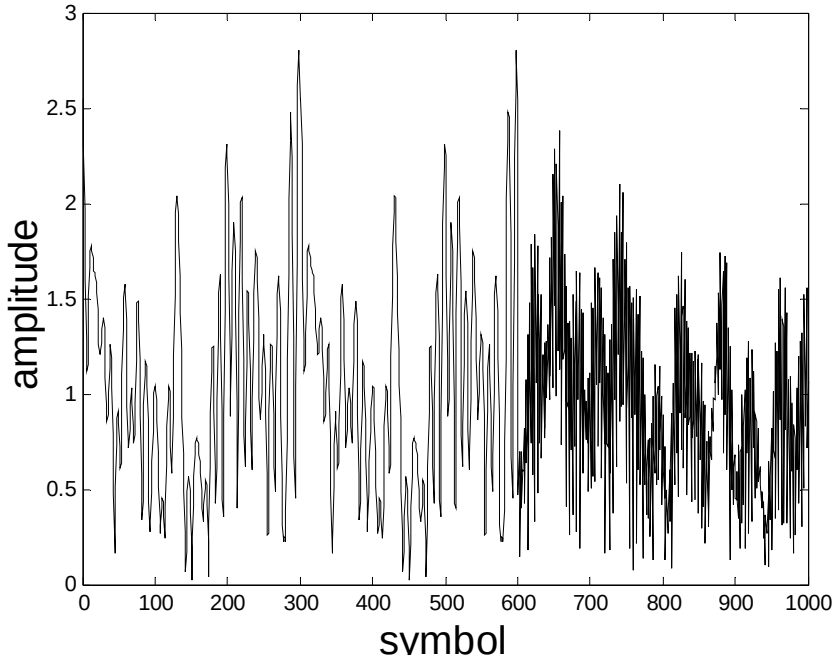


Figure 5.4 CSI for the physical model.

5.3 Long Range Prediction for AOFDM System

We now consider an OFDM system with K subcarriers, symbol (block) duration T_m, and adjacent subcarrier (tone) spacing Δf_s. Assume the channel bandwidth of the each subcarrier is much smaller than the coherence bandwidth and the channel state information does not change within one OFDM symbol duration $T = T_m+T_g$ but varies from symbol to symbol. The equivalent complex channel gain H[n, k] at n^{th} symbol block and k^{th} subcarrier can be modeled as the samples of the time-varying frequency selective channel in (5.6) with the time domain and frequency domain sampling interval T and Δf_s. From (2.11) and (2.12), the temporal and ensemble channel correlation functions for the OFDM symbols with block difference Δn and tone spacing Δk can be readily written as $R_T(\Delta nT, \Delta k\Delta f_s)$ and $R_E(\Delta nT, \Delta k\Delta f_s)$, respectively.

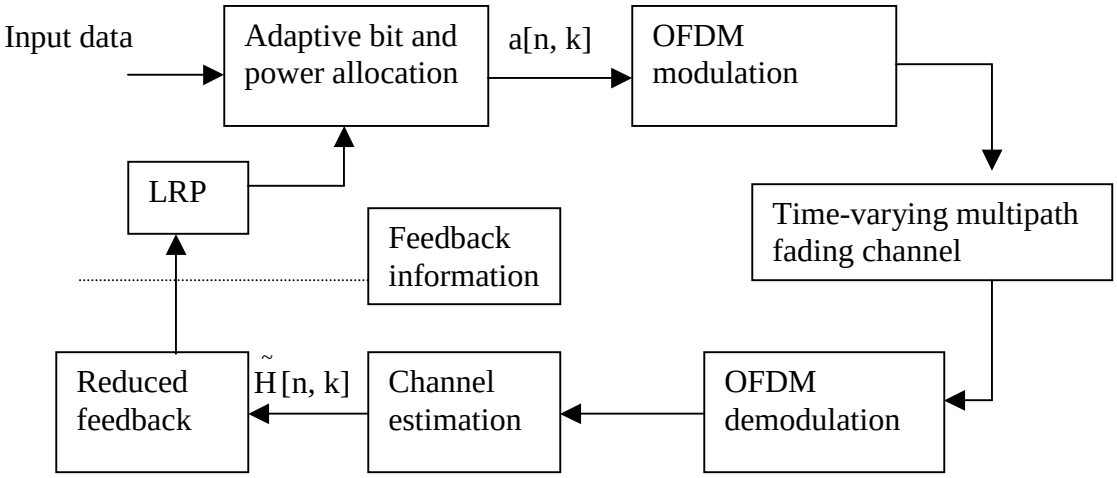


Figure 5.5 Block diagram of an adaptive OFDM system

The uncoded AOFDM system aided by the LRP and reduced feedback considered in this paper is depicted in Figure 5.5. The input data is allocated to the subcarriers according to the CSI fed back from the receiver. The LRP is employed at the transmitter to enhance the CSI accuracy. Let a[n, k] denote the complex baseband symbols at n^{th} block and k^{th} tone. The received signal after OFDM demodulation can be expressed:

$$X[n, k] = H[n, k]a[n, k] + w[n, k] \tag{5.11}$$

where w[n, k] is complex additive white Gaussian noise with variance $E[|w[n, k]|^2] = N_0$. Then frequency domain coherent channel estimation of the complex symbols associated with each of the K subcarriers is employed. A 2–D minimum mean square error channel estimator was proposed in [39]. Let

$$\tilde{H}[n, k] = H[n, k] + \tilde{w}[n, k] \tag{5.12}$$

denote the estimated CSI, where $\tilde{w}[n, k]$ is the estimation error modeled as white Gaussian noise with power spectrum $\tilde{N}_0$. We define the observation SNR (OSNR) = $E[|H[n, k]|^2]/\tilde{N}_0$. We normalized $E[|H[n, k]|^2$ to 1 throughout the chapter. Due to the correlated subcarriers, the estimated CSI $\tilde{H}[n, k]$, k=1…K, can be reduced and fed back to the long-range predictor at

the transmitter at low rate. Alternatively, the predictor can be placed at the receiver between the channel estimation and the reduced feedback blocks depending on the implementation issues such as complexity, performance and costs. The reduced feedback methods will be discussed in section 5.5. In this section, we assume that estimates of all subcarriers $\tilde{H}[n, k]$, k=1…K are available for the following long- range prediction algorithm.

5.3.1 MMSE Long Range Prediction

As we discussed in the chapter 2, the mobile fading channel over multiple carriers over a very short period of time can be characterized by a sum-of-sinusoids model as in (2.8) and (2.10) with the parameters $\{\underline{A},\underline{\theta},\underline{\tau},\underline{\phi}_1\}$. In this section, we derive the linear MMSE-based channel predictor for a multicarrier fading channel where $\{\underline{A},\underline{\theta},\underline{\tau}\}$ are fixed and the components of $\{\underline{\phi}_1\}$ are mutually independent random variables uniformly distributed on $[0,2\pi]$. One important parameter for the LRP is the sampling rate. For narrow band single carrier systems, the sampling rate of the LRP is much lower than the symbol rate (see chapter 4 or [16]). While the symbol interval in OFDM systems is longer, it is still beneficial to choose the sampling rate of the LRP lower than the symbol rate. Let $\tilde{H}[n, k]$ denote the estimated CSI (as in (5.12)) with a sample interval T_p (an integer multiple of the OFDM symbol interval T). The channel predictor for the CSI at the k^{th} tone and the n^{th} sample based on the p previously observed samples at K subcarriers can be constructed by:

$$\hat{H}[n, k] = \sum_{j=1}^{p} \sum_{m=1}^{K} d^{*}(j, m)\, \tilde{H}[n-j, m] \tag{5.13}$$

Provided that the correlation function (2.11) is known, the optimal filter coefficients $d_o(j, m)$ that minimize the conditional MSE $E_{\{\underline{\phi}_1\}}[|(H[n, k]-\hat{H}[n, k])|^2|\{\underline{A},\underline{\theta},\underline{\tau}\}]$ can be derived by using the orthogonaltiy principle [28]:

$$E_{\{\phi_1\}}[(H[n, k] - \hat{H}[n, k])^* \tilde{H}[n-g, l]|\{\underline{A},\underline{\theta},\underline{\tau}\}] = 0,\ g = 1\ldots p,\ l = 1\ldots K \qquad (5.14)$$

The resulting MMSE $J_{\min}$ is given by:

$$J_{\min} = 1 - \mathrm{tr}\,[\underline{\mathbf{D}}^H \underline{\mathbf{G}}]. \qquad (5.15)$$

where tr[•] is the trace of a matrix, the subscript H is *Hermitian* transpose and the matrix

$$\underline{\mathbf{D}} = \begin{bmatrix} d_o(1,1) & d_o(1,2) & \ldots. & d_o(1,K) \\ d_o(2,1) & . & . & . \\ . & . & . & . \\ . & . & . & . \\ d_o(p,1) & . & . & d_o(p,K) \end{bmatrix} \text{ and } \underline{\mathbf{G}} = \begin{bmatrix} r(1,k-1) & r(1,k-2) & \ldots. & r(1,k-K) \\ r(2,k-1) & . & . & . \\ . & . & . & . \\ . & . & . & . \\ r(p,k-1) & . & . & r(p,k-K) \end{bmatrix}$$

where $r(\Delta n, \Delta k) = R_T(\Delta n T_p, \Delta k \Delta f_s)$. This result serves as a theoretical foundation for our prediction problem and will be used in the performance analysis. The MMSE $J_{\min}$ depends on the scattering configuration $\{\underline{A},\underline{\theta},\underline{\tau}\}$ and number of waves N of the fading process. In general, our prediction method performs better when N is small. As N becomes large, the TACF (2.11) for the Jakes model and for our RPM (see sec. 2.4) approaches EACF (2.12) [35]. Therefore, $J_{\min}$ is upper-bounded by (5.15) with $r(\Delta n, \Delta k) \approx R_t(\Delta n T_p) R_f(\Delta k \Delta f_s)$. With this separation property, the closed-form expression for the optimal predictor and the MMSE given infinite past channel observation are derived in the appendix D. In Figure 5.6, we generate different scattering parameters $\{\underline{A},\underline{\theta},\underline{\tau}\}$ and calculate $J_{\min}$ for each experiment with k=2 (desired subcarrier) and K=3 (number of observation subcarriers) in (5.13). And Figure 5.7 shows the average of the calculated $J_{\min}$. It is observed that $J_{\min}$ calculated by TACF approaches that calculated by EACF when N is larger than 150. Moreover, the linear MMSE prediction results in the maximum likelihood (ML) prediction given past observations due to the assumption of joint Gaussian distribution of the observations and future samples [43].

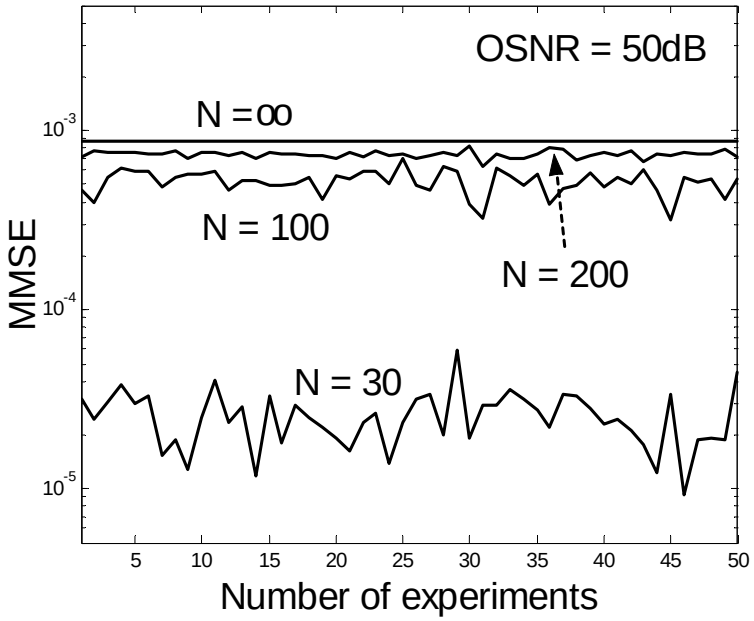


Figure 5.6 MMSE for the random phase model.

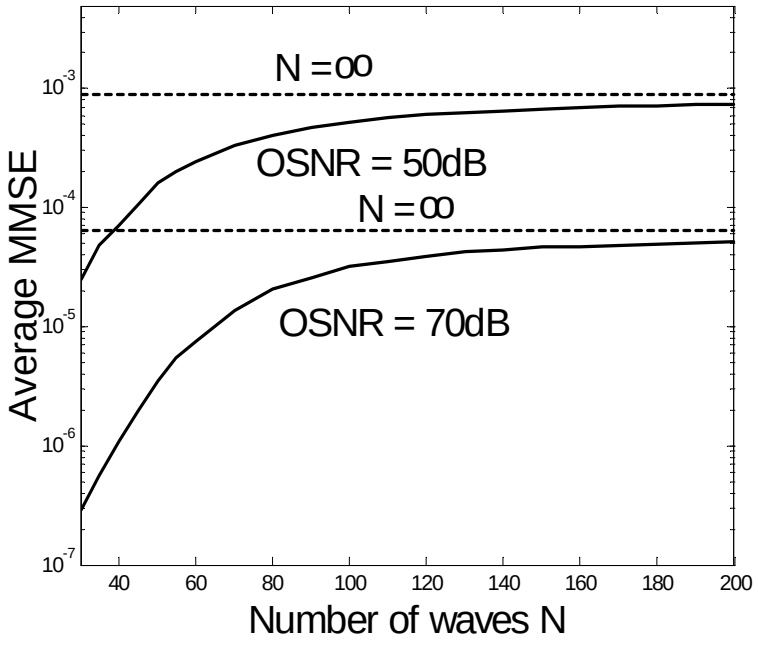


Figure 5.7 Average MMSE for the RPM with different N

5.3.2 Simplified LRP Method

In the linear prediction algorithm (5.13), the optimum MMSE is achieved by observing previous symbols of multiple subcarriers. However, this method is very complex in practice. Moreover, we have found that if the OSNR of the observed feedback samples is high, the improvement in the prediction accuracy when non-adjacent subcarriers' observations are used relative to utilizing just past samples of desired and adjacent subcarriers is negligible. This results from the fact that the CSI can be accurately predicted by the desired subcarrier for high OSNR and non--adjacent subcarriers is uncorrelated to the desired subcarrier due to large NFS. Actually, for the asymptotic case when the TAFC approaches the EACF factorization (2.12), and when the CSI is noiseless, it is sufficient to use just the past samples of the desired subcarrier to achieve the optimal MMSE performance (see Appendix D). Thus, we propose to simplify the algorithm by using only previously observed samples at subcarrier k to predict the CSI H[n,k]:

$$\hat{H}[n, k] = \sum_{j=1}^{p} d_j^*(n)\, \tilde{H}(n - j, k) \qquad k = 1,2,\ldots,K \tag{5.16}$$

Note that if the observation SNR is low, adjacent subcarriers can be easily incorporated to reduce the noise level at the cost of the system complexity. While in general, the coefficient vector $\underline{d}(n) = [d_1\ d_2 \ldots d_p]^T$ in (5.16) needs to be computed and adapted individually for each subcarrier, for our channel model, it is sufficient to employ the same filter coefficient vector $\underline{d}(n)$ to predict future CSI for each subcarrier. This method is justified by the fact that the flat fading coefficients of different subcarriers have approximately the same Doppler shifts (see (5.6)). It was shown in [18] that only the Doppler shifts associated with the scattering determine the prediction coefficients $\underline{d}(n)$ in (5.16). Hence the filter coefficient vector $\underline{d}(n)$ should remain tone-invariant resulting in significantly reduced computational complexity and

greatly improved tracking ability for the adaptive prediction methods discussed in the following sections since all feedback observations can be used jointly to update the coefficients. We call this method *simplified multiple carriers prediction* (SMCP). Note that this tone invariability can be generalized to the case when observation of several adjacent carriers are used provided that the same number of adjacent carriers is employed on each side of the desired carrier. This method extends to adaptive transmitter antenna diversity systems since the channels for all antennas have the same Doppler shifts [4].

The optimum MMSE channel prediction above relies on the knowledge of the time and frequency domain correlation functions (2.11). However, these correlation functions depend on the particular environment and usually are unknown. In addition, the coefficients $\underline{d}(n)$ in (5.16) needs to be computed adaptively as the Doppler shifts in (5.6) vary with time. In the following section, we employ the adaptive Least Mean Square (LMS) and Recursive Least Squares (RLS) algorithms, which do not require the knowledge of the correlation functions of the channel, to update the prediction filter coefficients for the OFDM system. The error between the desired response and the predicted CSI at subcarrier k is:

$$e[n, k] = H[n, k] - \sum_{j=1}^{p} d_j^*[n]\, \tilde{H}[n - j, k], \quad k = 1 \ldots K. \tag{5.17}$$

The average mean square error (AMSE) over all subcarriers is

$$\text{AMSE} = J(n) = \frac{1}{K}\sum_{k=1}^{K} |e[n, k]|^2 \tag{5.18}$$

This AMSE is used for updating the coefficients of the LMS and RLS algorithms. Note that the subcarriers are parameterized by the same $\{\underline{A},\underline{\theta},\underline{\tau}\}$ but different phases $\{\underline{\phi}_i\}$, where $\{\underline{\phi}_i\} = \{\phi_i(1) \ldots \phi_i(N), i = 1 \ldots K\}$. As the number of subcarriers K increases, the phases $\{\underline{\phi}_i\}$ become

randomized, and thus $J(n) \approx E_{\{\underline{\phi}_k\}}[e[n, k]^2|\{\underline{A},\underline{\theta},\underline{\tau}\}]$ for any k, and a lower bound on AMSE tends to the conditional MMSE for the single carrier LRP J_{smin}

$$J_{smin} = J_{min} \text{ for } K = 1 \tag{5.19}$$

where J_{min} is defined in (5.15). Note that using AMSE in SMCP, we adapt the coefficient vector $\underline{d}(n)$ jointly using the errors for all subcarriers. As discussed below, this improves accuracy and convergence rate relative to single carrier adaptive prediction [16,19,34]. We also observed that the prediction algorithm is more robust to noise in the feedback signals compared to the single carrier prediction [16,31] for both the LMS and RLS algorithms if the adjacent subcarriers are employed for prediction.

Assume the first p (filter length) samples are available for all subcarriers. Define the desired samples at time p+n as $\underline{\tilde{H}}_d[n]= [\tilde{H}[p+n, 1], \tilde{H}[p+n, 2],\ldots, \tilde{H}[p+n, K]]^T$ and the input p×p data matrix as $\underline{\tilde{\mathbf{H}}}[n] = \begin{bmatrix} \tilde{H}(p+n-1,1) & \tilde{H}(p+n-1,2) & \ldots. & \tilde{H}(p+n-1,K) \\ \tilde{H}(p+n-2,1) & . & . & . \\ . & . & . & . \\ . & . & . & . \\ \tilde{H}(n,1) & . & . & \tilde{H}(n,K) \end{bmatrix}$. This notation will be used in the following robust prediction algorithms.

5.3.3 LMS algorithm for AOFDM

The LMS algorithm belongs to the family of stochastic gradient algorithms and iteratively estimates the MMSE predictor filter [28]. The LMS algorithm uses the cost function $J(n)$ in (5.18). The equations for updating the predictor coefficients, shown in Table 5.1, are derived by calculating the gradient vector $\nabla J(n)$. From [28], the LMS learning curve depends on the step size μ and the eigenvalue spread of the correlation matrix, which can be

$\underline{d}_{px1}(0) = \underline{0}$

For each time instance, n=1,2…, compute

$\underline{\xi}_{px1}(n) = \underline{\tilde{H}}_d(n) - \underline{d}(n-1)^H \underline{\tilde{\mathbf{H}}}(n)$

$\underline{d}(n) = \underline{d}(n-1) + \mu \frac{1}{K} \underline{\tilde{\mathbf{H}}}(n) \underline{\xi}^*(n)$

Table 5.1 LMS algorithm for OFDM channel prediction

determined from the TAFC (2.11), of the input process. The selection of step size μ is a trade−off between the convergence rate and the excess MSE:

$$J_{ex}(n) = J(n) - J_{smin} \tag{5.20}$$

where $J(n)$ and J_{smin} are defined in (5.18) and (5.19), respectively. Similarly to the derivation for the single carrier prediction (see section 4.4), we derive the theoretical learning curve of the LMS for SMCP. Figure 5.8 demonstrate the comparison between the average theoretical learning curve and the numerical simulation for different scattering parameters $\{\underline{A},\underline{\theta},\underline{\tau}\}$. It is observed that while $J_{ex}(\infty)$ is larger for $\mu = 0.1$, it converges more rapidly than that for $\mu = 0.005$.

5.3.4 RLS algorithm for AOFDM

The RLS algorithm is a natural extension of the method of least squares. The idea of the algorithm is to design an adaptive filter such that given the least−squares estimate of the tap−weight vector at iteration n − 1, we may compute the updated estimate of the vector at iteration n upon the arrival of new data.

With the RLS algorithm, the predictor coefficients are calculated so that they minimize the error:

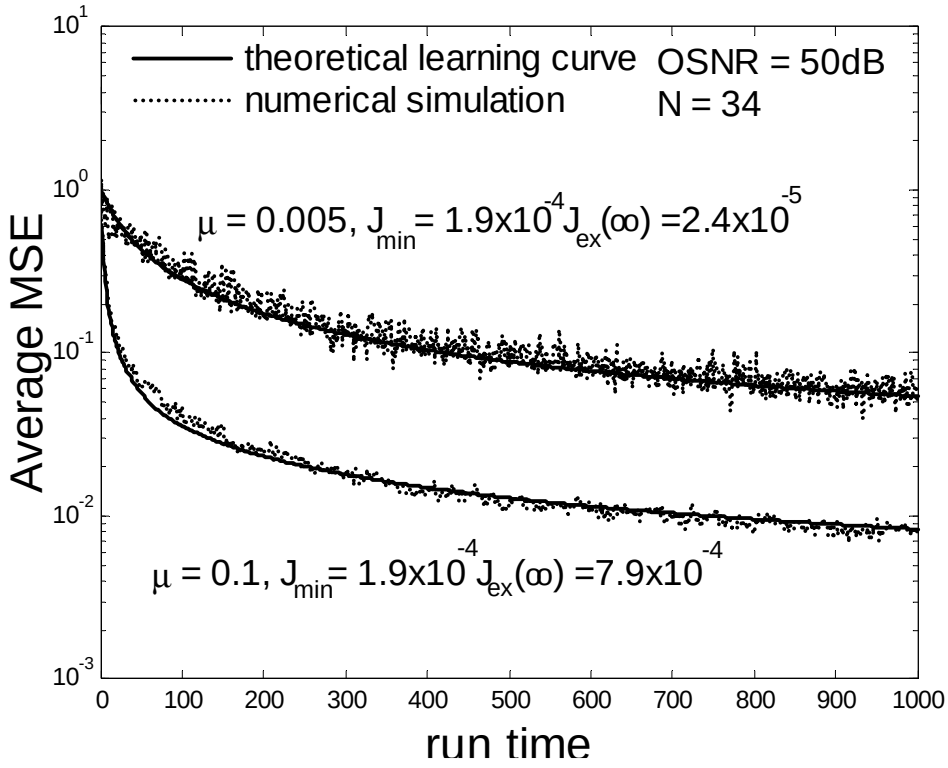


Figure 5.8 LMS learning curve. Normalized sample interval and prediction range in time $f_{dm}T = 0.2$. p = 50. K =100. $\Delta f_s = 0.005$.

$$\varepsilon[n] = \sum_{i=1}^{n} \sum_{k=1}^{K} \lambda^{n-i} |e[i, k]|^2 \tag{5.21}$$

where λ with $0 < \lambda \leq 1$ is the forgetting factor that accounts for possible non-stationary of the input CSI. The resulting update equation for the predictor coefficient vector $\underline{d}[n]$ is shown in Table 5.2. The parameter δ in the initialization procedure should be chosen small compared to the variance of the data samples. We use $\delta = 0.01$ in our simulation. For the RLS, the excess mean square error $J_{ex}(n)$ for SMCP is derived similarly to that for the single carrier case [28]. Its learning curve, unlike LMS algorithm, is independent of the eigenvalue spread of the input process and decays almost linearly with nK (the convergence rate is approximately K times faster than for single carrier prediction). For $\lambda = 1$, $J_{ex}(n)$ converges to

Initialize the algorithm by setting

$\underline{P}_{pxp}(0) = (\underline{\underline{\tilde{\mathbf{H}}}}(1)\underline{\underline{\tilde{\mathbf{H}}}}^{H}(1) + \delta\underline{I})^{-1}$

$\underline{d}_{px1}(0) = \underline{0}$

For each time instance, n=1,2…, compute

$$\underline{k}_{pxp}(n) = \frac{\lambda^{-1}\underline{P}(n-1)\underline{\underline{\tilde{\mathbf{H}}}}(n)}{1+\lambda^{-1}\underline{\underline{\tilde{\mathbf{H}}}}^{H}(n)\underline{P}(n-1)\underline{\underline{\tilde{\mathbf{H}}}}(n)}$$

$\underline{\xi}_{px1}(n) = \underline{\tilde{H}}_{d}(n) - \underline{d}(n-1)^{H}\underline{\underline{\tilde{\mathbf{H}}}}(n)$

$\underline{d}(n) = \underline{d}(n) + \underline{k}(n)\,\xi^{*}(n)$

$\underline{P}(n) = \lambda^{-1}\underline{P}(n-1) - \lambda^{-1}\underline{k}(n)\underline{\underline{\tilde{\mathbf{H}}}}^{H}(n)\underline{P}(n-1)$

Table 5.2 RLS algorithm for OFDM channel prediction

zero, and for λ close to 1 and large K, the $J_{ex}(\infty) \approx J_{smin}\frac{(1-\lambda)p}{2K}$. Hence for large K, $J_{ex}(\infty) \approx 0$. (While $J_{ex}(\infty) \geq 0$ in our example, for small values of K and proper choice of λ (close to 1 but not 1), it is possible for $J_{ex}(\infty)$ to be smaller than 0 due to the nonstationarity of the deterministic channel model (5.6) [48] and the fast tracking property of the RLS algorithm. Thus, SMCP improves the convergence rate and the steady state MSE for the RLS relative to the single carrier prediction [18,19,31]. Figure 5.9 demonstrates the learning curve of the RLS algorithm. It is observed that the RLS algorithm can converge rapidly to the J_{smin} for large K (=100). In addition, when K =10, and λ = 0.9, the AMSE can be smaller than J_{smin}. This simulation result reveals that the performance of the AMSE can be improved by properly choosing the number of K and the forgetting factor λ for SMCP RLS algorithm.

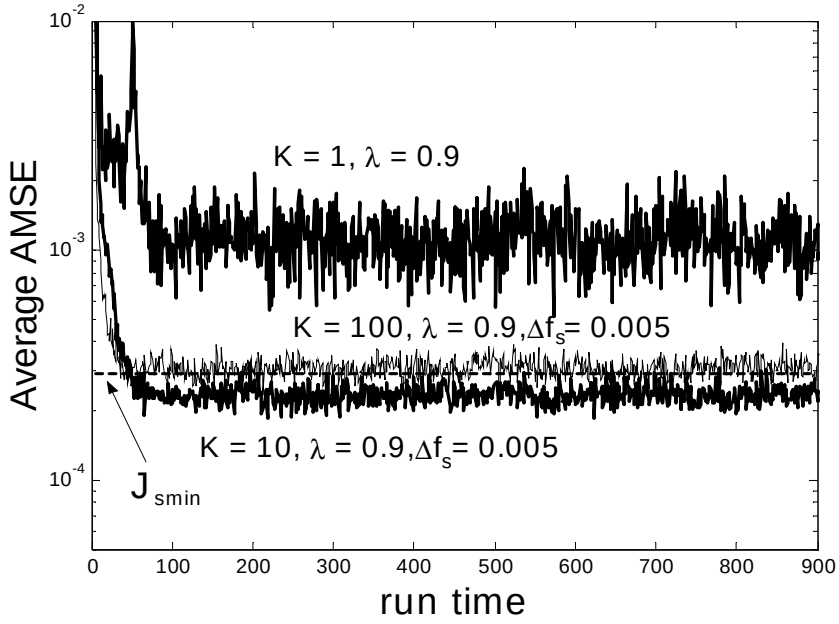


Figure 5.9 RLS learning curve. Normalized sample interval and prediction range in time $f_{dm}T_p$ = 0.2. p = 50.

5.3.5 Numerical simulations

We use the RPM and the physical model (see section 5.1.3) to validate the performance of the LRP for the OFDM system. To test the performance of our prediction algorithm on the fading channel modeled by the RPM, N = 34 is chosen and multiple deterministic channel realizations are generated by using independent angles $\{\underline{\theta}\}$ and propagation delays $\{\underline{\tau}\}$. We apply the LRP for each independent realization and calculate the AMSE (5.18). Furthermore, the J_{smin} (5.19) averaged over these independent realizations is presented for comparison. Note that when we use the RPM, each channel realization has different TACF. Thus, we can test the tracking ability and the prediction accuracy (AMSE) of our prediction algorithm in different scattering environments. This task cannot be accomplished using the Jakes model, where the incident angles are fixed. The prediction filter length p in (5.16) is 50. The maximum Doppler shift of 100 Hz is used in both models.

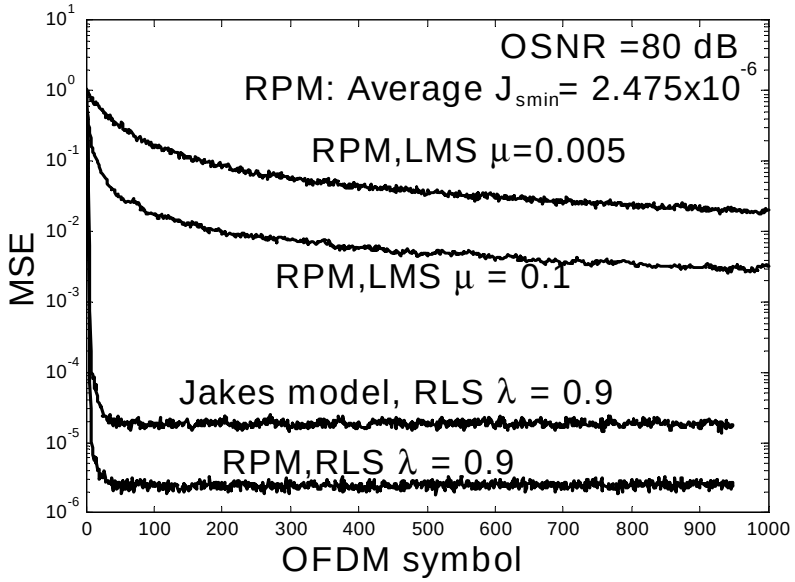


Figure 5.10 Performance of different adaptive prediction methods for the random phase model.

The *rms* delay spreads is approximately 1μs in both channel models. To construct an OFDM symbol, assume that the entire channel bandwidth, 800kHz, is divided into 128 subcarriers. The symbol duration is 160μs. An additional 5μs guard interval is used to provide protection from ISI due to channel multipath delay spread. Thus the total block length is 165μs and the subcarrier symbol rate is approximately 6KHz. For each subcarrier, the fading signal is sampled at the low rate of 466Hz for the LRP (the prediction range is 1/466Hz ≈ 2ms). In this thesis, we assume reliable channel estimation and high effective SNR (80 dB) of the observed CSI. While actual SNR of the observed samples might be much lower, it is possible to employ noise reduction techniques due to very low sampling rate of the LRP [16,19,27]. Interpolation is utilized to predict channel coefficients at the subcarrier symbol rate [16,19,31].

Figure 5.10 demonstrates the average AMSE (5.18) over all the channel realizations for the SMCP method for the RPM model. When these results are compared with the single

carrier prediction, we find that the excess mean square error for the LMS algorithm $J_{ex}(n)$ (5.20) of these two approaches is approximately the same given the same step size μ.. This result suggests that the number of the subcarriers employed for the SMCP is insignificant for determining the learning curve of the LMS algorithm once the step size μ is chosen. The MSE curve shown for μ=0.005 corresponds to both methods. However, the single carrier algorithm diverges for large μ, while for the SMCP (5.16), μ can be chosen as large as 0.1 without divergence, thus improving the convergence rate. As the NFS increases, larger step size μ can be chosen, resulting in faster convergence. This phenomenon results from the fact that weak correlation between the subcarriers due to large NFS results in small bias in the estimate of the gradient vector $\nabla J(n)$. While the RLS has higher computational complexity than the LMS algorithm, its learning curve and the excess MSE $J_{ex}(n)$ (5.20) are significantly improved relative to the LMS. It is observed that the RLS algorithm converges rapidly with almost no excess MSE for $\lambda = 0.9$, whereas the LMS algorithm converges more slowly with significant excess MSE relative to the RLS algorithm. We also demonstrate the RLS algorithm for the Jakes model. We observe that for $\lambda = 0.9$, the AMSE of the more realistic RPM is better than that of the Jakes model.

In Figure 5.11, the SMCP is explored for the physical model. It is demonstrated that the RLS algorithm is good at tracking the non-stationary channel. During the transition period (from samples 500 to 700 in Figure 5.4), the forgetting factor $\lambda = 0.1$ has better tracking ability than $\lambda = 0.9$. Hence it is more robust to the non-stationary environment. The tracking results for the LMS algorithm are much poorer with a relatively high MSE during and even after the transition period.

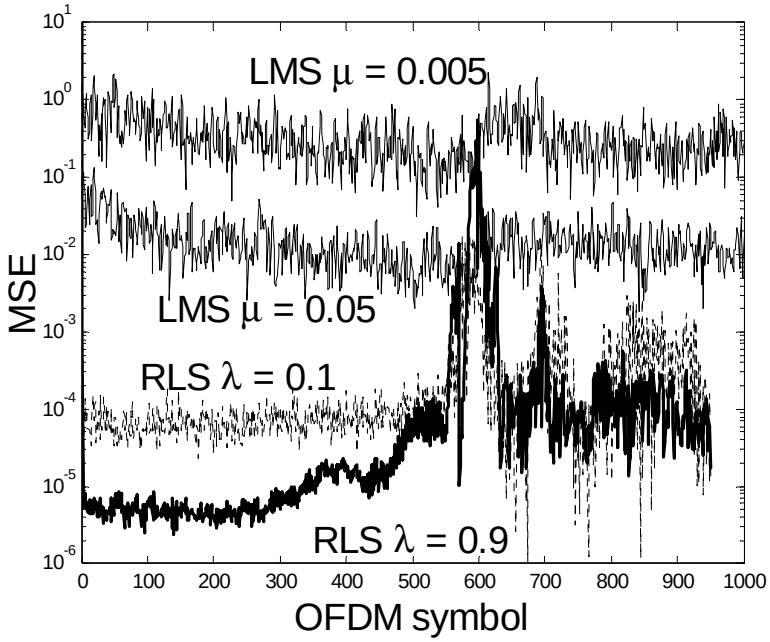


Figure 5.11 Performance of different adaptive prediction methods for the physical model

5.4 Adaptive Bit and Power Loading Algorithm

Adaptive OFDM system, similarly to adaptive modulation for single carrier flat fading channel [16,21,22,24,31,57], involves optimizing the modulation level and the transmit power over the entire frequency band to maximize the spectral efficiency. In a frequency selective fading channel, some subcarriers experience a deep fade while others are subject to channel gain. The spectral efficiency can be improved by allocating more bits to those subcarriers with favorable channel conditions than to those in a deep fade, as motivated by the "*water filling"* distribution [45]. Several practical integer-bit and power allocation algorithms have been addressed in [10,11,26,36] that perform the optimum or near-optimum loading of bit and power in an OFDM frame. However, these methods perform adaptive bit and power allocation assuming perfect CSI. In this section, we employ channel loading optimization in the presence of imperfect CSI that results from prediction errors under the *bit*

rate maximization (BRM) criterion, where the goal is to allocate the limited energy among the subcarriers to maximize the overall bit rate subject to a target bit error rate constraint [11]. A simplified loading method similar to [10] is compared with the optimal Hughes-Hartogs algorithm [26].

5.4.1 Robust adaptive and power allocation

For each subcarrier we employ rectangular M(i)–QAM modulation [12] where M(1)=0, M(i)=2^{i-1}, i = 2…6. Let $\hat{c}$ denote the CSI obtained from the linear prediction algorithm (5.16) and c the actual complex gain at a certain subcarrier. Hence $\hat{c}$ and c are jointly complex Gaussian and their amplitudes $\hat{\alpha}$ and α are both Rayleigh distributed. For subcarrier k, let P_k (= $E[|a[n, k]|^2]$ (5.11)) denote the transmitted signal power of the complex M(i)–QAM symbol that is determined by allocation algorithm. (Note the sum of the allocated powers for all subcarriers does not exceed the total power constraint P_{total}). Assume each subcarrier has the same noise power N_0 (see (5.11)). The SNR $\gamma(i)$ (= P_k/N_0) required to employ M(i)–QAM modulation based on the predicted channel gain $\hat{\alpha}_k$ at the k^{th} subcarrier can be found by numerical search to solve $P_e^*(M(i),\gamma(i),\hat{\alpha}_k) = BER_{tg}$, where P_e^* is demonstrated in (3.18). Note P_e^* depends on the correlation coefficient ρ between $|\alpha_k|^2$ and $|\hat{\alpha}_k|^2$. Once $\gamma(i)$ are calculated for each modulation level and each subcarrier, they are used to implement the Hughes-Hartogs and the simplified algorithm. The only difference in the implementation (relative to the perfect CSI case) is that the SNR $\gamma(i)$ is used in place of the ideal SNR required to achieve the BER with M(i)-QAM [12].

The optimal Hughes-Hartogs algorithm is based on a relatively simple scheme called *successive bit allocation algorithm*. In each iteration, the algorithm assigns one bit to the

subcarrier requiring the least additional power to reliably transmit that bit. The additional power per bit required by the modulation with M(i+1) levels relative to the modulation with M(i) levels can be calculated by:

$$\gamma_b = \frac{\gamma(i+1) - \gamma(i)}{\log_2 M(i+1) - \log_2 M(i)} \tag{5.22}$$

The parameter γ_b is computed for each subcarrier in each iteration and the bits and power are designated to the subcarrier with the minimum γ_b. The algorithm stops when no additional bits can be added without violating the total power constraint.

While the Hughes-Hartogs algorithm is optimal and simple to implement, it requires intensive sorting and computation and thus renders the algorithm impractical for applications where the number of subcarriers are large. We employ a simplified sub-optimal algorithm in [10]. This algorithm has two steps. The first step is called *equal power allocation*. The same amount of energy P_{total}/K is assigned to each subcarrier at subcarrier k for selecting the initial modulation level:

$$M = \max \{M(i) \mid P_e^*(M(i), P_{total}/(KN_0), \hat{\alpha}_k) < BER_{tg}\}, \; k = 1 \ldots K \tag{5.23}$$

After the modulation level is decided, the power distribution P_k at subcarrier k is calculated so that $P_e^*(M, P_k/N_0, \hat{\alpha}_k) = BER_{tg}$. This power allocation step exploits the fact that the difference between the optimal water-filling distribution and the flat-energy distribution is minimal. Note the total power usage must be less than the total power constraint P_{total}. By applying the *successive bit allocation algorithm,* the second step performs adaptive bit allocation using the rest of the power.

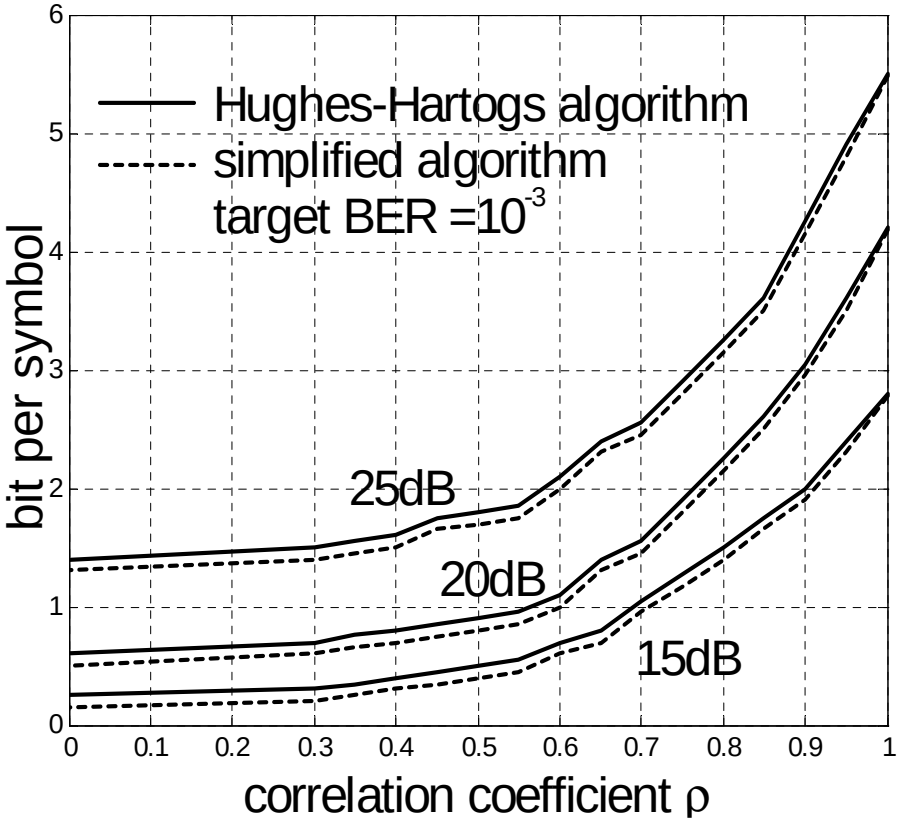


Figure 5.12 BPS vs. ρ for different SNR for adaptive OFDM system. SNR = P_{total}/KN_0.

5.4.2 Numerical simulation

The performance of the aforementioned AOFDM algorithm depends on the correlation coefficient ρ between $|\alpha_k|^2$ and $|\hat{\alpha}_k|^2$, k = 1…K. The average bits per symbol (BPS) vs. the correlation coefficient ρ vs. the average SNR constraint given by $\frac{P_{total}}{KN_0}$ is shown in Figure 5.12, where we assume each subcarrier has the same prediction accuracy ρ. The BER constraint for each subcarrier is 10^{-3}. The correlation $\rho = 1$ corresponds to perfect prediction, while $\rho = 0$ represents the worst case when the BPS of the adaptive modulation converges to that of the non-adaptive M-QAM for given SNR and bit error rate constraint

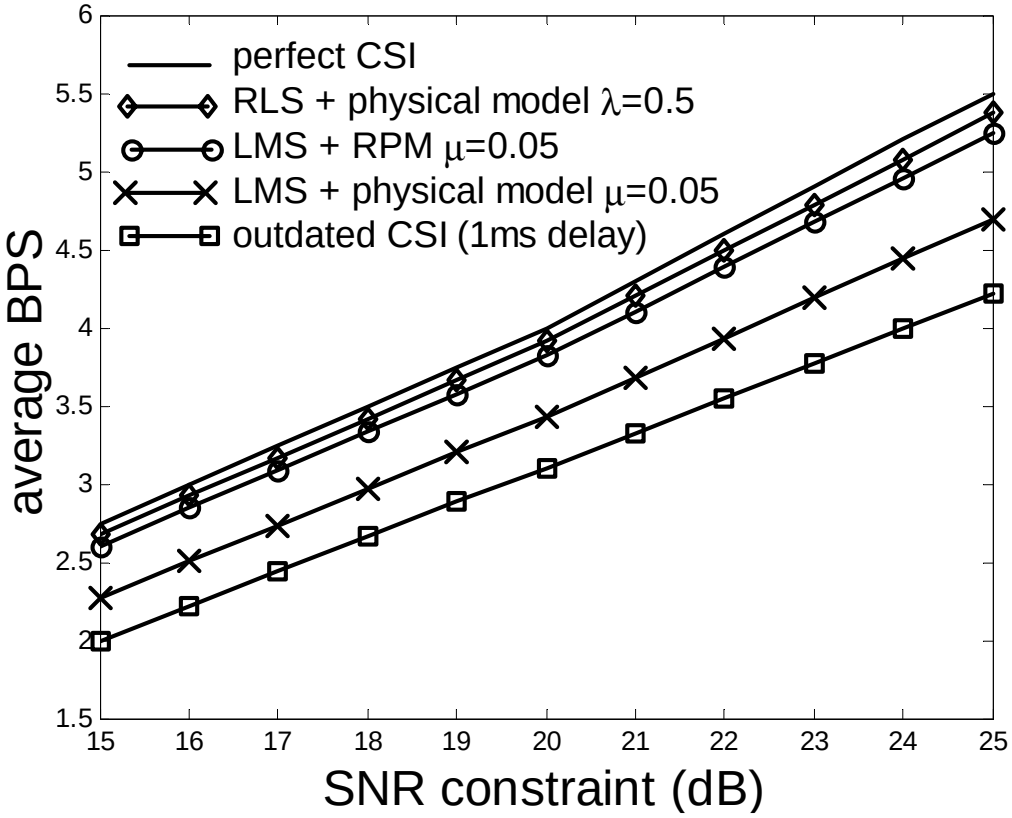


Figure 5.13 Comparison of average BPS performance for adaptive OFDM aided by different prediction methods for the RPM and physical model.

BER_c. It is observed that the simplified algorithm is near-optimal when ρ is close to 1 and has performance loss less than 0.1 BPS for ρ <<1 compared with the optimal Hughes-Hartogs algorithm.

We use the RPM and the physical model to validate the performance of our AOFDM algorithm aided by the LRP. Perfect feedback is assumed. The target BER for the adaptive OFDM system is 10^{-3}. The system parameters are described in section 5.2.5. The average BPS of the AOFDM for different prediction algorithms over the RPM and physical channel models is plotted in Figure 5.13. Note that the spectral efficiency will be slightly less than the average BPS due to the guard interval used to eliminate the ISI. Comparison reveals that the

RLS has better performance than the LMS algorithm for the RPM and non-stationary physical model. The performance of the RLS algorithm for the RPM is almost near-optimal, whereas the loss is only about half dB for the physical model compared to the perfect knowledge of CSI. The performance of the AOFDM using the outdated CSI for the RPM is also shown in Figure 5.13. To alleviate the mismatch of the delayed and future CSI, an approach to calculate thresholds based on the delayed CSI was studied in [24]. We found that even very small delay causes significant loss of the bit rate for fast vehicle speeds when accurate LRP is not utilized. For example, the delay of 1ms for f_{dm} =100Hz corresponds to correlation coefficient $\rho = 0.9$ in (3.24). This results in the bit rate loss about 0.5 bits/symbol for the target BER = 10^{-3} and the SNR per symbol = 15 dB assuming stationary Rayleigh fading channel, while prediction with RLS allows to achieve the bit rate of about 2.6 bits/symbol for the non-stationary, and the near-ideal rate for the stationary model. Thus, accurate LRP is required to achieve the bit rate gain of adaptive OFDM system for fast vehicle speeds and realistic delays.

5.4.3 Block Loading in Frequency

To reduce signaling overhead complexity and to accommodate the hardware constraints, we investigated the performance of block adaptive loading in frequency domain, where neighboring subcarriers are grouped into a sub-channel and use the same modulation level and power control. In addition, block loading can reduce the complexity of the loading algorithm since the number of sub-channels is fewer than the number of subcarriers. Similarly to the block loading method in 3.5, we employ two different power control methods. They are per-subcarrier (PS) and per-sub-channel (PC) power control. In the PS method, power can be adjusted for each subcarrier in that sub-channel, and in the PC method,

power remains constant for the whole sub-channel. Next, we discuss the Hughes-Hartogs block loading algorithm.

Let L be the number of subcarriers in a sub-channel and $\{\alpha_1,\alpha_2,\ldots\alpha_L\}$ be the amplitudes of CSI at subcarrier 1 to subcarrier L in a sub-channel during an OFDM symbol interval. Assume perfect CSI and threshold T_i, i = 1…4 is decided as in (3.11). For that sub-channel, the required power to employ M(i)-QAM for the PS and PC methods are $S_f(i) = \frac{1}{L}\sum_{k=0}^{L-1} \bar{S}(\frac{T_i}{\alpha_k})^2$ and $S_f(i) = \{\bar{S}\beta \mid \frac{1}{L}\sum_{k=0}^{L-1} P_e(M_i, \bar{\gamma}\beta\ (\alpha_k)^2) = BER_{tg}\}$, respectively. For the Hughes-Hartogs algorithm, The additional power per bit required by the modulation with M(i+1) levels relative to the modulation with M(i) levels for each sub-channel is calculated by:

$$\gamma_b = \frac{S_f(i+1)-S_f(i)}{\log_2 M(i+1)-\log_2 M(i)} \tag{5.24}$$

The parameter γ_b is computed for each sub-channel in each iteration and the bits and power are designated to the sub-channel with the minimum γ_b. The algorithm stops when no additional bits can be added without violating the total power constraint. Figure 5.14 demonstrates the numerical results for the block loading over frequency domain. The total number of subcarrier K =160, and the normalized subcarrier spacing $\Delta f_s\sigma = 0.005$. It is observed that performance loss is negligible if $\Delta f_s L\sigma < 0.1$ (L = 20). Thus, 8 sub-channels, relative to 160 subcarriers, are created for the loading algorithm, resulting in significant complexity reduction. Furthermore, block loading method in time described in 3.5 can be incorporated resulting in block loading in time and frequency adaptive loading method. Note the simplified loading algorithm can also be easily accommodated for the block loading to further reduce the computation complexity.

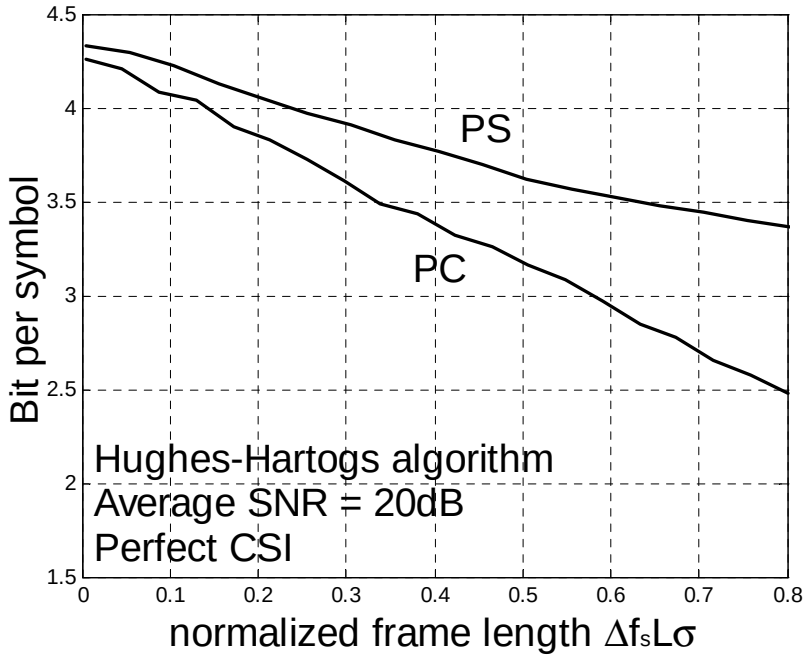


Figure 5.14 Bit per symbol for block loading adaptive modulation and power control.

5.5 Feedback Load Reduction

In addition to adaptive channel loading, many other adaptive transmission techniques can be implemented in a wireless OFDM system. These include adaptive coding, adaptive transmitter antenna diversity and interference suppression at the transmitter. In a closed−loop system the transmitter obtains the CSI for this adaptive transmission from the receiver via a feedback channel. The prediction can be performed either at the transmitter or at the receiver, depending on the complexity and feedback requirements. In some adaptive transmission applications, it might be feasible to predict at the receiver and feed back the CSI parameters derived from the predictions (e.g., the modulation level and power for adaptive modulation). In other adaptive transmission techniques (e.g., adaptive antenna diversity), predicted channel gains are often required at the transmitter, and thus channel gains have to be fed

back, independent of whether the prediction is accomplished at the transmitter or receiver. This CSI is required for all subcarriers. It is desirable to minimize this feedback load since it consumes resources that would otherwise be used for data. At the same time, the feedback signal should carry enough information so the transmitter can perform reliable adaptive transmission.

In this section, we explore several methods for reducing the feedback of the OFDM signal vectors $\underline{\tilde{H}} = [\tilde{H}(n, 1) \ldots \tilde{H}(n, K)]$ (the estimates in (5.12) sampled at low rate) while insuring accurate reconstruction at the transmitter. Since $\underline{\tilde{H}}$ is modeled as Gaussian, the estimates of the reconstructed signals are formed as linear combinations of the signals that are fed back. The performance is measured by the correlation coefficient ρ (3.24) between the reconstructed signals and actual CSI for each subcarrier and is dependent on the feedback density (FD) given by:

$$\text{FD} = \frac{\text{Number of fed back symbols}}{\text{Total number of subcarriers}} \tag{5.25}$$

We also define the normalized feedback density (NFD) as:

$$\text{NFD} = \frac{\text{FD}}{\text{normalized subcarrier frequency separation}} \tag{5.26}$$

We assume non-quantized channel observations. In practice, the fedback symbols need to be quantized. However, due to the low feedback rate (sampling rate) required for the LRP and the reduced feedback feasible for AOFDM, the quantization level can be chosen large without significantly reducing the achievable bit rate. Thus, the actual performance is expected to be close to that obtained for ideal non-quantized channel observations.

5.5.1 Karhunen-Loeve (KL) Low Rank Modeling

The K-L method [28] requires the knowledge of the eigen-vectors of the correlation matrix of the feedback signal vector at both the transmitter and the receiver. For the fading channel characterized as WSSUS with the frequency domain correlation function $R_f(\Delta f)$ (2.12), the K×K correlation matrix defined as $\mathbf{R} = E[\tilde{\underline{H}}\,\tilde{\underline{H}}^H]$ is computed from $R_f(\Delta f)$. Let $\underline{q}_1, \underline{q}_2 \ldots \underline{q}_K$ be the eigenvectors associated with K eigenvalues ($\lambda_1 > \lambda_2 > \ldots \lambda_K$) of the matrix $\mathbf{R}$. The vector $\tilde{\underline{H}}$ can be expressed as a linear combination of these eigenvectors as

$$\tilde{\underline{H}} = \sum_{i=1}^{K} v_i \underline{q}_i. \tag{5.27}$$

The coefficients v_i of the expansion are zero-mean, uncorrelated random variables defined by the inner product $v_i = \underline{q}_i^H \tilde{\underline{H}}$. From (5.27), we can reduce the feedback by sending only v_i for i = 1…m and to approximately reconstruct $\tilde{\underline{H}}$ at the transmitter by $\hat{\underline{H}} = \sum_{i=1}^{m} v_i \underline{q}_i$, m<K. Hence the feedback density (5.25) is m/K. The reconstruction error vector is defined by $\underline{e} = \hat{\underline{H}} - \tilde{\underline{H}}$. It can be shown that the mean square error is $E[\underline{e}^H \underline{e}] = \sum_{i=m+1}^{K} \lambda_i$, [28], Thus, accurate reconstruction is achieved if the eigenvalues $\lambda_{m+1}, \ldots, \lambda_K$ are very small. The number of significant eigenvalues depends on the rms delay spread, frequency separation, and the number of the subcarriers and can be shown to be small for typical OFDM channels.

5.5.2 IDFT Method and Direct Feedback Method

Since computation of the basis of the *K-L* low rank modeling [24] requires the knowledge of the channel correlation function, we propose to utilize the discrete Fourier

basis and transform the CSI $\underline{\tilde{H}}$ using the inverse discrete Fourier transform (IDFT). This choice is meaningful since the IDFT corresponds to the channel impulse response. The K-point IDFT of the CSI $\underline{\tilde{H}}$ is given by

$$I(m) = \frac{1}{K}\sum_{k=1}^{K} \tilde{H}(n, k) exp\{j2\pi \frac{k-1}{K}(m-1)\},\ m = 1 \ldots K \tag{5.28}$$

The samples I(m) are relatively small for $m > \tau_{max}\Delta f K$, where τ_{max} is the maximum excess delay, Δf is the subcarrier frequency separation and K is the total number of subcarriers. This suggests that $FD > \tau_{max}\Delta f$ is required to obtain good performance. In OFDM channels, K is chosen much larger than $\tau_{max}\Delta f K$ (i.e. $\Delta f << 1/\tau_{max}$) to avoid intersymbol interference (ISI). Thus, the transformed signal I(m) can be truncated, fed back to the transmitter and reconstructed by the DFT.

Alternatively, we can directly feed a subset of the CSI samples $\underline{\tilde{H}}$ back to the transmitter without any transformation. The feedback signals are sampled uniformly over the entire frequency band. The original signal can be reconstructed by interpolation or using the MMSE criterion if the correlation functions are known. This method results in different accuracy for each subcarrier and increases the complexity of the bit and power allocation. We call this method *direct reduced feedback.*

For all reduced feedback methods, the performance depends on the *rms* delay spread and the subcarrier frequency separation. Furthermore, it will affect the LRP accuracy and hence the performance of the adaptive loading algorithm. Therefore, there is a tradeoff between the feedback load and the performance of the AOFDM.

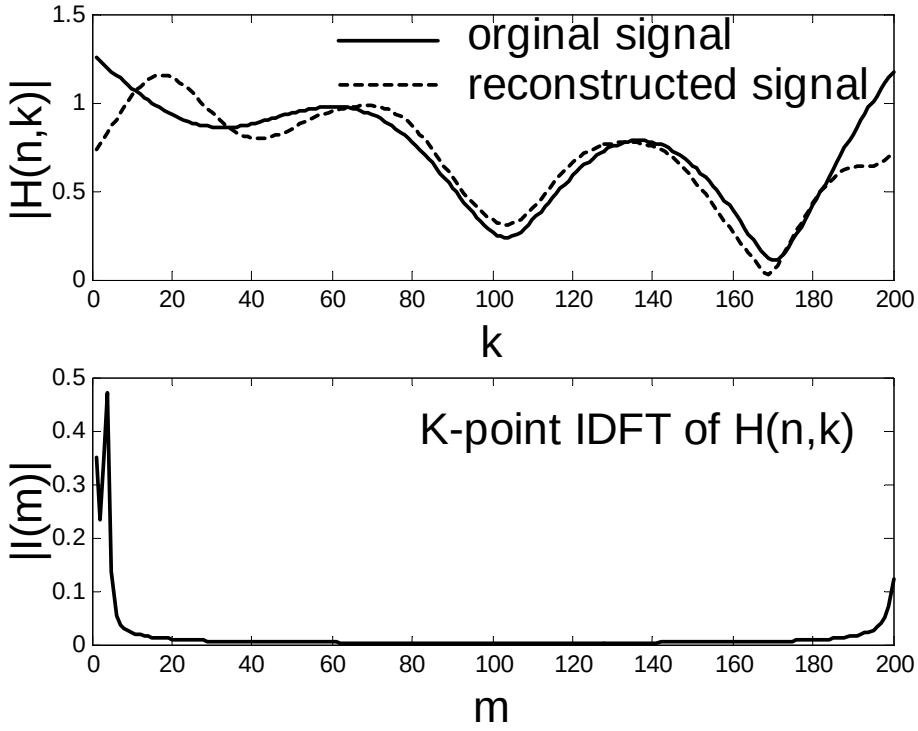


Figure 5.15 Reduced feedback using IDFT

5.5.3 Numerical Simulation for AOFDM and Reduced Feedback

Figure 5.15 shows the deterministic channel response, its K-point IDFT (K=200), and the reconstructed signal (5 symbols are fed back). Note that significant "anti-causal" part of |I(m)|, m=180…200, is observed since the observation signal $\tilde{H}(n, k)$ is only a truncation of the channel frequency response. Discarding this anti-causal part will degrade the performance of the IDFT method. The performance loss can be mitigated by increasing the total number of subcarriers K to reduce the anti-causal part.

In Figure 5.16, we compare the performance of the KL and IDFT reduced feedback methods. The normalized subcarrier frequency separation is 0.005, and the FD (5.25) of 0.025 is used. This implies that only 5 complex symbols out of 200 are fed back to the transmitter. The SNR of the observed signal $\underline{\tilde{H}}$ is 80dB. The correlation between the

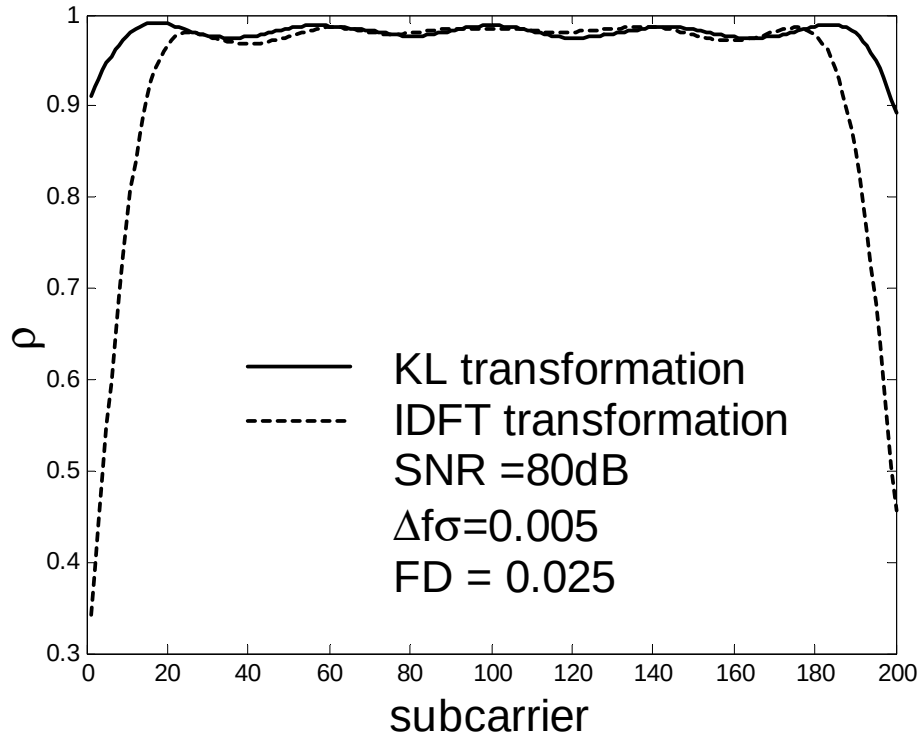


Figure 5.16 Performance comparisons of the reduced feedback methods.

reconstructed signal and the original signal ρ (3.24) is computed for each subcarrier. It is observed that the parameter ρ is close to 1 for the center subcarriers and drops sharply for the IDFT method for the edge subcarriers**.** The performance can be improved by increasing the total number of subcarriers K to reduce the anti-causal part of I(m) for the same FD. In practical systems, the edge subcarriers can be used as guard bands [39] and do not require feedback.

The performance of the prediction algorithm using reduced feedback methods for different SNR constraint is shown in Figure 5.17 for the random phase channel model (see 2.4.1). Assume the OSNR of the estimated signal at the receiver is 80dB. The signals are reduced by the methods described in section 5.5 and fed back to the transmitter. The transmitter predicts the next OFDM symbol based on the received signal, estimates the

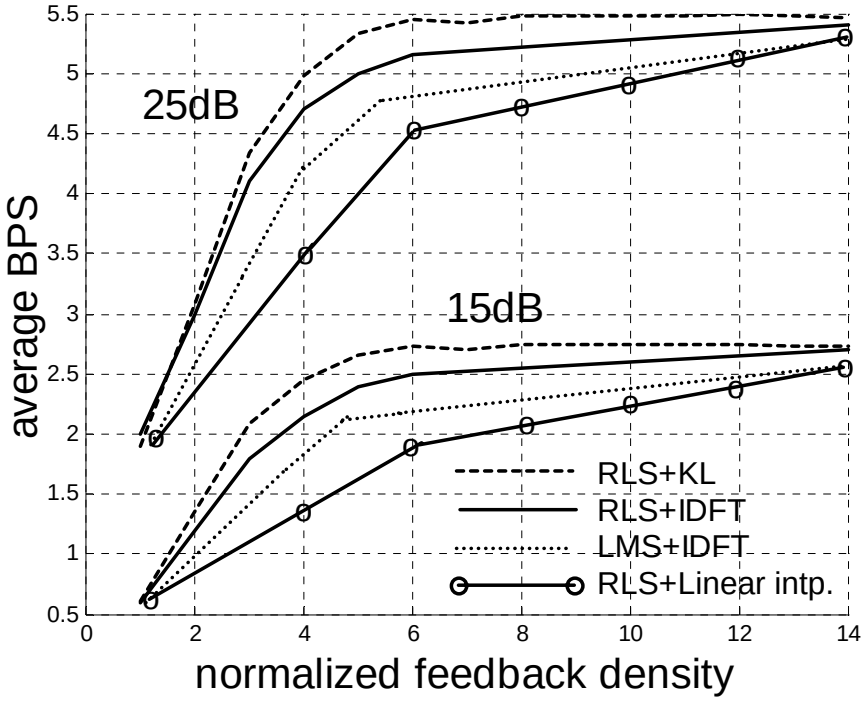


Figure 5.17 Performance comparison for different reduced feedback methods.

prediction accuracy ρ, and performs adaptive bit and power allocation. It shows that for the KL method, performance is near optimal when the NFD (5.26) is larger than 6. This implies from (5.26) that only 4 symbols need to be fed back for a 128-subcarriers OFDM system with normalized subcarrier frequency separation 0.005. For the IDFT method, the performance loses 0.5 BPS at NFD = 6 as opposed to 1 BPS loss for the direct reduced feedback method with linear interpolation. The selection of the feedback density is a trade off between the feedback load and the prediction accuracy, and hence the transmission rate.

CHAPTER 6

CONCLUSIONS

This thesis has concentrated on realistic fading channel modeling, long range channel prediction for wireless mobile fading channel and adaptive modulation with power control. In particular, we investigated adaptive modulation aided by observations of another carrier and adaptive channel loading for the wireless OFDM system.

A commonly used continuous-time mobile fading channel model, which is a weighted sum of complex sinusoids, is employed for our study. We derived the temporal average correlation function and ensemble average correlation function over time and frequency domain. These correlation functions serve as the theoretical foundation for the study of the long range channel prediction. In addition to this parameter-fixed sum-of-sinusoids model, a novel physical model that can generate realistic non-stationary fading process is introduced.

A discrete rate adaptive modulation method employing M–QAM and continuous power control is investigated to maximize the transmission throughput over wireless fading channels. We compared its performance with the algorithm optimized for the assumed channel statistics and observed that our practical and simple method results in less than 0.5 dB power loss relative to the optimal *discrete rate continuous power* method over Rayleigh fading channel. Block loading methods when the same modulation method and power control are employed in time domain are investigated for practical constraints. Similar block loading algorithms can be applied in frequency domain over a large number of subcarrier. Imperfect CSI is also considered in the design of reliable adaptive modulation methods. It is found that small delay causes significant loss of the bit rate for adaptive transmission systems for fast vehicle speeds. Therefore, to realize the potential of adaptive transmission methods, the

channel variations have to be predicted at least several milliseconds, or tens to hundreds of data symbols, ahead.

We studied linear LRP for the OFDM system and for the single carrier system based on the observation of another carrier. One important feature of the LRP is its low sampling rate (on the order twice the maximum Doppler shift), which results in large memory span for a fixed filter length. The low sampling rate also reduces the feedback rate.

For the single carrier prediction system, we concentrate on the scenario where we observe a received uplink signal at the carrier frequency f^1 and attempt to predict the downlink signal at the carrier frequency f^2 without feedback from the mobile. We derive the MMSE LRP that utilizes the time and frequency domain correlation function of the Rayleigh fading channel. An adaptive MMSE prediction method is also proposed. We demonstrated that significant bit rate gains can be achieved relative to non-adaptive systems for realistic channel parameters, and that increased frequency separation and multipath delay limit the performance of adaptive transmission. We also use the physical model to identify a typical case where the *rms* delay is approximately constant and a challenging case where *rms* delay is severely and rapidly changes. We test the limitation of the adaptation rate to the variation of the *rms* delay. The results give valuable insights into designing adaptive transmission methods for correlated carriers and multicarrier systems.

For the multicarrier OFDM system, we investigated the long range prediction based on past channel observations over multiple subcarriers. We found that while larger number of observed subcarriers can result in better prediction performance, the improvement in the prediction accuracy when non-adjacent subcarriers' observations are used relative to utilizing just past samples of desired and adjacent subcarriers is negligible. We develop a simplified

multiple carriers prediction method (SMCP) based on the fact that fading for all subcarriers can be approximately characterized by the same auto-regressive model. Adaptive prediction methods based on the LMS and RLS algorithms for SMCP are developed. For the LMS algorithm, the learning curve depends on the step size and the eigen-structure of the input CSI. Larger step size relative to that for the single carrier LMS algorithm can be employed for SMCP resulting in faster convergence. While the RLS algorithm is more complex than the LMS algorithm, it has significantly improved learning curve and the excess MSE and is capable of tracking non-stationary signals. Numerical simulations reveal that, for the physical and the standard sum-of-sinusoids model, the performance of the adaptive OFDM based on the CSI predicted by the RLS algorithm approaches that based on the perfect CSI. An optimal and a simplified adaptive bit and power loading algorithm are investigated assuming imperfect CSI, and their performance is compared. It is observed that the simplified algorithm is near-optimal with accurate CSI and has performance loss less than 0.1 BPS with unreliable CSI compared with the optimal Hughes-Hartogs algorithm. Finally, several methods were developed to reduce the feedback load, and it was shown that the IDFT method offers significant feedback load reduction while maintaining near-optimal spectral efficiency.

APPENDIX A

DERIVATION OF THE ENSEMBLE AVERAGE CORRELATION FUNCTION

We derive the ensemble correlation function in (2.12). The CSI at carrier frequency f^1 and f^2can be characterized by (2.8) and (2.10). The correlation function with time difference Δt and frequency separation Δf can be calculated as $E[c(t, f)\, c^*(t+\Delta t, f+\Delta f)] =$

$$E[\sum_{n=1}^{N} A(n) exp\{j(2\pi f_d(n)\, t - \phi(n))\} \sum_{m=1}^{N} A(m) exp\{j(-2\pi f_d(m)\,(t + \Delta t) + \phi(m) + 2\pi \Delta f \tau(m))\}] \quad (A.1)$$

Assume uncorrelated scattering, i.e., $E[A(n)A(m)] = 0$ for $n \neq m$, (A.1) can be simplified as

$$E[\sum_{n=1}^{N} A(n)^2\, exp\{j(2\pi f_d(n)\, t - \phi(n))\}\, exp\{j(-2\pi f_d(n)\,(t + \Delta t) + \phi(n) + 2\pi \Delta f \tau(n))\}]$$

$$= E[\sum_{n=1}^{N} A(n)^2 \exp\{-j2\pi f_d \Delta t\} \exp\{j2\pi \Delta f \tau(n)\}] \quad (A.2)$$

It is reasonable to assume the amplitude, Doppler shift and delay spread are independent random variables and hence the correlation function can be separated as:

$$\sum_{n=1}^{N} E[A(n)^2]\, E[\exp\{-j2\pi f_d(n) \Delta t\}]\, E[\exp\{j2\pi \Delta f \tau(n)\}] \quad (A.3)$$

$$= \Omega\, R_t(\Delta t)\, R_f(\Delta f)$$

where Ω is the total average received power from all multipath components, and $R_t(\Delta t) = E[\exp\{-j2\pi f_d(n) \Delta t\}]$, and $R_f(\Delta f) = E[\exp\{j2\pi \Delta f \tau(n)\}]$. Evaluation of the expectation in (A.3) requires that we specify the probability density function for the angle of incident θ of the arrival wave and for the delay spread τ. By assuming isotropic channel, i.e. θ are uniformly distributed around $[0,2\pi]$, $R_t(\Delta t)$ can be calculated as [35]

$$E[\exp\{-j2\pi f_d \Delta t\}] = \frac{1}{2\pi}\int_0^{2\pi} \exp\{-j2\pi f dm\cos(\theta)\Delta t\} d\theta = J_0(2\pi f dm \Delta t) \qquad (A.4)$$

where J_0 is the zero-order Bessel function of the first kind. Likewise, by assuming the delay spread are exponentially distributed [35] as in (2.13), $R_f(\Delta f)$ can be evaluated as

$$E[\exp\{j2\pi\Delta f\tau_n(t)\}] = \int_0^{2\pi} \exp\{j2\pi\Delta f\tau)\frac{1}{\sigma} exp\{-\tau/\sigma\} d\tau$$

$$= \frac{1}{1+(2\pi\Delta f\sigma)^2} + j\frac{2\pi\Delta f\sigma}{1+(2\pi\Delta f\sigma)^2} \qquad (A.5)$$

APPENDIX B

CORRELATION COEFFICIENT ρ

We derive the formula for the correlation coefficient ρ in (3.25) and (4.10). As we discussed in this chapter 3, the performance of the AM with imperfect CSI significantly depends on the correlation coefficient ρ between $\alpha = |c|^2$ and $\hat{\alpha} = |\hat{c}|^2$, where c is the actual CSI and $\hat{c}$ is the estimate of c or related to c. The definition of the correlation coefficient ρ is shown in (3.24).

B.1 Proof of (3.25)

The covariance between α^2 and $\hat{\alpha}^2$ can be calculated as

$$\mathrm{cov}(\alpha^2, \hat{\alpha}^2) = E[((x_1^2+y_1^2) - \Omega_1)((x_2^2+y_2^2) - \Omega_2)] =$$

$$E[x_1^2x_2^2+x_1^2y_2^2+y_1^2x_2^2+y_1^2y_2^2-\Omega_2(x_1^2+y_1^2)- \Omega_1(x_2^2+y_2^2)+ \Omega_1\Omega_2] \quad \text{(B.1)}$$

It is shown in [43, p 160] that if the random variables x and y are jointly normal with zero mean, then

$$E[x^2y^2] = E[x^2]E[y^2] + 2E^2[xy] \quad \text{(B.2)}$$

Since we assume x_1, x_2, y_1 y_2 are jointly normal, (B.1) can be simplified as

$$\mathrm{cov}(\alpha^2, \hat{\alpha}^2) = 4(u_I + u_Q) \quad \text{(B.3)}$$

For α being Rayleigh distribution, it is shown in [43, p 111]

$$\mathrm{Var}(\alpha_1^2) = E[\alpha_1^4] - E^2[\alpha_1^2] = 2\Omega_1^2 - \Omega_1^2 = \Omega_1^2 \quad \text{(B.4)}$$

$$\mathrm{Var}(\alpha_2^2) = \Omega_2^2$$

Substituting (B.2) and (B.3) for (3.24) gives

$$\rho = \frac{\mathrm{Cov}(\alpha^2,\hat{\alpha}^2)}{\sqrt{\mathrm{Var}(\alpha^2)\mathrm{Var}(\hat{\alpha}^2)}} = 4\frac{u_I^2 + u_Q^2}{\Omega_1\Omega_2} \quad \text{(B.5)}$$

B.2 Proof of (4.10)

For c_2 being the MMSE estimate of c_1, from orthogonality principle, $E[(c_1 - c_2)^* c_2] = 0$, the power Ω_2 can be expressed as

$$\Omega_2 = E[c_1^* c_2] = 2u_I + j2u_Q \tag{B.6}$$

Since Ω_2 is a real number, $u_Q = 0$ and $\Omega_2 = 2u_I$. From (B.4), Ω_2 can be written as $\Omega_2 = \rho\Omega_1$, Therefore, the MMSE can be expressed as:

$$\text{MMSE} = \Omega_1 - \Omega_2 = \Omega_1(1 - \rho) \tag{B.7}$$

APPENDIX C

MMSE LRP FOR INFINITE NUMBER OF PAST CHANNEL OBSERVATIONS

We derive the *discrete time* linear MMSE shown in (4.11) for one step prediction given infinite number of past observations. From (4.3), the one-step prediction at frequency f^2 given infinite past channel observations at f^1 is given by:

$$\hat{c}(f^2,n) = \sum_{j=1}^{\infty} d_j^* \tilde{c}(f^1,n-j) \qquad \text{(C.1)}$$

The filter coefficient d(n) can be determined from the orthogonality principle [28] as:

$$\sum_{j=1}^{\infty} d(j) r_n(k-n) = g(k) \qquad k=1,2\ldots\infty \qquad \text{(C.2)}$$

where $r_n(j-k) = E[\tilde{c}(f^1,k)\tilde{c}^*(f^1,j)] = R_t((j-k)\Delta t$/at sub-sampling rate fs) $+ N_0\delta(j-k)$ is the discrete time autocorrelation of the noisy fading channel samples, and $g(k) = E[\tilde{c}(f^1,n-k)c^*(f^2,n)] = R_f(\Delta f)R_t(k\Delta t$/at sub-sampling rate fs). Since $r_n(n)$ is a correlation sequence, we can represent $r_n(n)$ as:

$$r_n(n) = \sum_{j=-\infty}^{0} r_n^-(j)\, r_n^+(n-j) \qquad \text{for all n} \qquad \text{(C.3)}$$

where $r_n^-(n)$ and $r_n^+(n)$ are sequences with the properties that $r_n^-(n) = 0$ when $n>0$ and $r_n^+(n) = 0$ when $n<0$. From [53], d(z), the Z–transform of the filter d(n), can be derived as

$$d(z) = R_f(\Delta f)\left[z - \frac{1}{r_z^+(z)}\, r_n^+(0)z\right] \qquad \text{(C.4)}$$

where $r_z^+(\bullet)$ is the Z–transform of $r^+(k)$. With the knowledge of filter coefficient, we can express the $MMSE_{p=\infty} = E[|c(n) - \hat{c}|^2]$ as:

$$MMSE_{p=\infty} = R_n(0) - \sum_{n=1}^{\infty} d^*(n)g(n) = R_n(0) - b_0 \tag{C.5}$$

where $R_n(0) = r_n(0) - N_0$ and b_0 is the 0^{th} term of the convolution of $d^*(-n)$ and g(n). Let B(z) = $g(z)Z[d^*(-n)] = g(z)\, d(z^{-1})$, where Z [d(–n)] is the z transform of d (–n). Then B(z) can be expressed as

$$B(z) = [1 - \frac{1}{r_z^-(z)} r_n^+(0)]\, [r_z(z) - N_0] \tag{C.6}$$

Hence b_0 can be decided as

$$b_0 = [r_n(0) - N_0 - r_n^+(0)^2 + N_0] \tag{C.7}$$

Then we have

$$MMSE_{p=\infty} = R_n(0) - b_0 = R_n(0) - [r_n(0) - r_n^+(0)^2] = r_n^+(0)^2 - N_0 \tag{C.8}$$

where $r_n^+(0)^2$ can be decided as [53]:

$$r_n^+(0)^2 = \exp\{\frac{1}{2\pi}\int_{-\pi}^{\pi} \ln[r_w(w)]dw\} = \exp\{\frac{1}{2\pi}\int_{-\pi}^{\pi} \ln[R_w(w)+N_0]dw\} \tag{C.9}$$

where $R_w(w) = \sum_{n=-\infty}^{\infty} R_n(n)\, e^{-jwn}$ is the folded power spectrum of the c(n) and $R_n(n) = E[C(k)$ C*(k+n)] is the autocorrelation function of the C (n).

APPENDIX D

MMSE LRP FOR OFDM SYSTEMS

It is discussed in the chapter 5 that J_{min} is upper-bounded by (5.15) if the correlation can be separated by the time and frequency correlation functions, i.e., $r(\Delta n, \Delta k) \approx R_t(\Delta n T_p) R_f(\Delta k \Delta f_s)$. Based on this separation property, we derive the closed form solution of the optimal predictor $d_o(j,m)$ in (5.13), and the asymptotical MMSE for the infinite past channel observations. In addition, we also prove that when the CSI is noiseless, it is sufficient to use just the past samples of the desired subcarrier to achieve the optimal MMSE performance.

D.1 LRP of OFDM

From the separation property, the correlation function can be expressed as:

$$E[\tilde{H}^*[j, m]\, \tilde{H}[g, l] = r_t[j - g] r_f[m - l] + N_0 \delta[j - g] \delta[m - l] \quad \text{(D.1)}$$

The optimal channel predictor for the CSI at the k^{th} tone and the n^{th} block based on the p previous observed samples and K subcarriers can be can be derived by orthogonaltiy principle as in (5.14). By applying (D.1), we can rewrite (5.14) as:

$$r_t[g]\, r_f[k - l] = \sum_{j=1}^{p} \sum_{m=1}^{K} d(j, m)\, r_t[g - j]\, r_f[m - l] + N_0\, d(g, l),\ g = 1 \ldots p,\ l = 1 \ldots K \quad \text{(D.2)}$$

Denote $\underline{r}_f = \begin{bmatrix} r_f[k-1] \\ r_f[k-2] \\ . \\ . \\ r_f[k-K] \end{bmatrix}$, $\underline{d}_f(j) = \begin{bmatrix} d(j,1) \\ d(j,2) \\ . \\ . \\ d(j,K) \end{bmatrix}$, and $\underline{R}_f = \begin{bmatrix} r_f[0] & r_f[1] & \ldots. & r_f[K-1] \\ r_f^*[1] & . & . & . \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ r_f^*[K-1] & . & . & r_f[0] \end{bmatrix}$

Eq. (D.2) can be expressed in vector form as:

$$r_t[g]\,\underline{r}_f = \sum_{j=1}^{p} r_t[g-j]\,\underline{R}_f\,\underline{d}_f(j) + N_0\,\underline{d}_f(g),\ g = 1\ldots p \tag{D.3}$$

Let the eigendecomposition of $\underline{R}_f$ be

$$\underline{R}_f = \underline{U}^H\,\underline{Q}\,\underline{U} \tag{D.4}$$

where $\underline{U}$ is a unitary matrix and $\underline{Q}$ is a diagonal matrix with diagonal element q_l. Substituting (D.4) for (D.3) gives

$$r_t[g]\,\underline{Q}^{-1}\,\underline{U}\,\underline{r}_f = \sum_{j=1}^{p} r_t[g-j]\,\underline{U}\,\underline{d}_f(j) + N_0\,\underline{Q}^{-1}\,\underline{U}\,\underline{d}_f(g) \tag{D.5}$$

Denote $\tilde{\underline{r}}_f = \underline{Q}^{-1}\,\underline{U}\,\underline{r}_f$, and $\tilde{\underline{d}}_f(j) = \underline{U}\,\underline{d}_f(j)$, then (D.5) can be rewritten as

$$r_t[g]\,\tilde{\underline{r}}_f = \sum_{j=1}^{p} r_t[g-j]\,\tilde{\underline{d}}_f(j) + N_0\,\underline{Q}^{-1}\,\tilde{\underline{d}}_f(g) \tag{D.6}$$

Equivalently, (D.6) can be written as:

$$r_t[g]\,\tilde{r}_f[l] = \sum_{j=1}^{p} r_t[g-j]\,\tilde{d}(j, l) + \frac{N_0}{q_l}\,\tilde{d}(g, l),\ l = 1\ldots K,\ g = 1\ldots p \tag{D.7}$$

where $\tilde{r}_f[l]$, $\tilde{d}(j, l)$ are the l^{th} elements of $\tilde{\underline{r}}_f$ and $\tilde{\underline{d}}_f(j)$, respectively. And q_l is the l^{th} diagonal element of $\underline{Q}$. Denote

$$\underline{r}_t = \begin{bmatrix} r_t[1] \\ r_t[2] \\ . \\ . \\ r_t[p] \end{bmatrix},\ \tilde{\underline{d}}_t(l) = \begin{bmatrix} \tilde{d}(1, l) \\ \tilde{d}(2, l) \\ . \\ . \\ \tilde{d}(p, l) \end{bmatrix},\ \underline{R}_{t,\,l} = \begin{bmatrix} r_t[0]+\frac{N_0}{d_l} & r_t[1] & \ldots. & r_t[p-1] \\ r_t[1] & . & . & . \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ r_t[p-1] & . & . & r_t[0]+\frac{N_0}{d_l} \end{bmatrix}$$

From (D.7), the filter coefficient $\tilde{\underline{d}}_t(l)$ can be solved by:

$$\underline{\tilde{d}}_t(l) = \tilde{r}_f[l]\ \underline{R}_{t,\,l}^{-1}\ \underline{r}_t \qquad \text{(D.8)}$$

Let $\underline{\tilde{D}}_{Kxp} = \begin{bmatrix} \underline{\tilde{d}}_t^{\,T}(1) \\ \underline{\tilde{d}}_t^{\,T}(2) \\ . \\ . \\ \underline{\tilde{d}}_t^{\,T}(K) \end{bmatrix}$. Equivalently, $\underline{\tilde{D}}_{Kxp}$ can be expressed as $\underline{\tilde{D}}_{Kxp} =$

$\begin{bmatrix} \underline{\tilde{d}}_f(1) & \underline{\tilde{d}}_f(2) & ... & \underline{\tilde{d}}_f(p) \end{bmatrix}$. The optimal coefficient d(j,m) can be found as

$$\underline{D}_{Kxp} = \begin{bmatrix} d(1,1) & d(2,1) & \ldots. & d(p,1) \\ d(1,2) & . & . & . \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ d(1,K) & . & . & d(p,K) \end{bmatrix} = \underline{U}^H\ \underline{\tilde{D}} \qquad \text{(D.9)}$$

The average power of the predicted symbol $E[|\hat{H}[n, k]|^2]$ is required to calculate the MMSE.

From (5.13), $E[|\hat{H}[n,k]|^2] = E[\sum_{j_1=1}^{p}\sum_{m_1=1}^{K} d^*(j_1,m_1)\tilde{H}[n-j_1,m_1]\sum_{j_2=1}^{p}\sum_{m_2=1}^{K} d(j_2, m_2)\ \tilde{H}^*[n-j_2, m_2]]$

$= \sum_{j_1=1}^{p}\sum_{m_1=1}^{K} d^*(j_1, m_1)\sum_{j_2=1}^{p}\sum_{m_2=1}^{K} d(j_2, m_2)\ E[\tilde{H}^*[n-j_2, m_2]\ \tilde{H}[n-j_1, m_1]]$

$= \sum_{j_1=1}^{p}\sum_{m_1=1}^{K} d^*(j_1, m_1)\ r_t[j_1]\ r_f[k - m_1] = tr[\underline{D}^H\underline{G}]$,

where $\underline{G} = \begin{bmatrix} r_t(1)r_f(k-1) & r_t(2)r_f(k-1) & \ldots. & r_t(p)r_f(k-1) \\ r_t(1)r_f(k-2) & . & . & . \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ r_t(1)r_f(k-K) & . & . & r_t(p)r_f(k-K) \end{bmatrix}$

Hence, the MMSE = $E[|(H[n, k]-\hat{H}[n, k])|^2] = 1 - tr\ [\underline{D}^H\ \underline{G}]$.

D.2. MMSE for Infinite Past Observation

If the parameter p goes to infinity, (D.7) can be rewritten as:

$$r_t[g] = \sum_{j=1}^{\infty} r_t[g-j] \frac{\tilde{d}(j, l)}{\tilde{r}_f[l]} + \frac{N_0}{q_l} \frac{\tilde{d}(g, l)}{\tilde{r}_f[l]}, \; l = 1 \ldots K, \; g = 1 \ldots p \qquad (D.10)$$

Notice the solution $\frac{\tilde{d}(j, l)}{\tilde{r}_f[l]}$ in (D.10) is equivalent to that solved in the single carrier prediction investigated in the Appendix C with observation noise power $\frac{N_0}{q_l}$. Therefore,

$$\sum_{j=1}^{\infty} \tilde{d}^*(j, l)\, r_t[j] = b_l\, \tilde{r}_f[l], \; l = 1 \ldots K \qquad (D.11)$$

where $b_l = r_t[0] + \frac{N_0}{q_l} - r_l^+[0] = 1 + \frac{N_0}{q_l} - r_l^+[0]$, $r_l^+[0] = \exp\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln[R_w(w) + \frac{N_0}{q_l}] dw \}$, and $R_w(w) = \sum_{n=-\infty}^{\infty} r_t(n)\, e^{-jwn}$

Eq. (D.11) can be expressed in vector form as:

$$\sum_{j=1}^{\infty} r_t[j]\, \underline{\tilde{d}}_f(j)^H = [\, b_1 \tilde{r}_f[1] \quad b_2 \tilde{r}_f[2] \ldots b_K \tilde{r}_f[K] \,] = \underline{v}^T \qquad (D.12)$$

where $\underline{v}^T = \sum_{j=1}^{\infty} r_t[j]\, \underline{d}_f(j)^H \underline{U}^H$. To calculate the MMSE, the average power of the predicted symbol $E[|\hat{H}[n, k]|^2]$ can be expressed as $E[|\hat{H}[n, k]|^2] = \sum_{j=1}^{\infty} \sum_{m=1}^{K} d^*(j, m)\, r_t[j]\, r_f[k-m]$, and it can be written in vector form as $E[|\hat{H}[n, k]|^2] = \sum_{j=1}^{\infty} r_t[j]\, \underline{d}_f(j)^H \underline{r}_f = \underline{v}^T \underline{U}\, \underline{r}_f$. Hence

$$\text{MMSE} = E[|(H[n, k] - \hat{H}[n, k])|^2] = 1 - \underline{v}^T \underline{U}\, \underline{r}_f \qquad (D.13)$$

D.3. Noiseless Observation

In this section, we prove that when the CSI is noiseless, it is sufficient to use just the past samples of the desired subcarrier to achieve the optimal MMSE performance. For noiseless observation, the correlation function, (D.1) can be simplified as

$$E[\tilde{H}^*[j, m]\,\tilde{H}[g, l] = r_t[j - g]r_f[m - l] \quad (D.14)$$

Denote H[n, k] the desired channel information. The observation can be decomposed as $\tilde{H}$ [n–j, m] = {$\tilde{H}$ [n–j, m] – E[•]$\tilde{H}$ [n–j, k]} + E[•]$\tilde{H}$ [n–j, k]

= $\tilde{H}_o$ [n–j, m] + E[•]$\tilde{H}$ [n–j, k], j = 1…p, m = 1…K.

where E[•] = E[$\tilde{H}$ [n–j, m] $\tilde{H}^*$ [n–j, k]] = $r_t[0]r_f[k - m] = r_f[k - m]$. It can be proved readily from (D.14) that $\tilde{H}_o$ [n–j, m], j=1…p is orthogonal to $\tilde{H}$ [n–g, k], g = 1…p and to desired symbol H[n, k], i.e.,

$$E[\tilde{H}_o\,[n–j, m]\,\tilde{H}^*\,[n–g, k]] = 0,\ j= 1…p,\ g = 1…p,\ m = 1…K \quad (D.15)$$

$$E[\tilde{H}_o\,[n–j, m]\,H^*\,[n, k]] = 0,\ j = 1…p,\ m = 1…K \quad (D.16)$$

Thus, from (5.13), the predicted symbol $\hat{H}$ [n, k] can be expressed as $\hat{H}[n, k] = \tilde{H} + \tilde{H}_o$, where $\tilde{H}$ is spanned by $\tilde{H}$ [n–j, k], j=1…p, and $\tilde{H}_o$ is spanned by $\tilde{H}_o$ [n–j, m], j=1…p, m = 1…K. Notice that

$$E[\tilde{H}_o\tilde{H}^*] = 0. \quad (D.17)$$

The MSE = $E[|(H[n, k]–\hat{H}[n, k])|^2] = E[|(H[n, k]– \tilde{H})|^2] + |\tilde{H}_o|^2$. The second equality is based on (D.15–D.17). Consequently, to minimize the MSE, it is best to restrict attention to $\tilde{H}$ spanned by $\tilde{H}$[n–j, k], j = 1…p.

BIBLIOGRAPHY

[1] F. Adachi, K. Ohno, A. Higashi, T. Dohi, and Y. Okumura, "Coherent Multicode DS-CDMA Mobile Access," *IEICE Transaction on Commun.*, Vol. E79-B, pp. 1316-1324, Sep. 1996.

[2] M. S. Alouini and A. J. Goldsmith, "Area Spectral Efficiency of Cellular Mobile Radio Systems," *IEEE Transactions on Vehicular Technology*, vol. 48, July 1999, pp. 1047 - 1066.

[3] J. B. Anderson, J. Jensen, S. Holdt Jensen, F. Frederiksen, "Prediction of Future Fading Based on Past Measurements," *IEEE Signal Processing Mag.*, vol. 17, No. 3, May 2000, pp. 62-75.

[4] A. Arredondo, K.R. Dandekar, Guanghan Xu, "Vector Channel Modeling and Prediction for the Improvement of Downlink Received Power," *IEEE Trans. Commun*, vol. 50, no. 7, pp. 1121-1129, July 2002.

[5] J. A. C. Bingham, "Multicarrier Modulation for Data Transmission: An idea whose time has come," *IEEE commun. Mag.*, pp. 5-14, May 1990.

[6] H. L. Bertone, *Radio Propagation for Modern Wireless Systems*, Prentice-Hall, 2000.

[7] E. Biglier, G. Caire, and G. Taricco, "Coding and Modulation under Power Constraints," *IEEE Personal Commun*. Vol. 5, June 1998, pp. 32-39.

[8] R. W. Chang, "Synthesis of Band Limited Orthogonal Signals for Multichannel Data Transmission," *Bell Syst, Tech. J.*, Vol. 45, pp. 1775-196, Dec. 1996.

[9] J. K. Cavers, "An Analysis of Pilot Symbol Assisted Modulation for Rayleigh Fading Channels," *IEEE Journal on Selected Areas in Communications*, vol. 7, pp. 1347-1355, Dec. 1989.

[10] A. Czylwik, "Adaptive OFDM for Wideband Radio Channels," *Proc. IEEE GLOBECOM'96*, Nov. 1996, pp. 713-718.

[11] J. Campello, "Practical Bit Loading for DMT," *Proc. IEEE ICC '99*, Vol. 2, 6-10 June, 1999.

[12] K. Cho, D. Yoon, "On the General BER Expression of One- and Two-Dimensional Amplitude Modulations," *IEEE Trans. Commun.*, vol. 50, pp. 1074-1080, July 2002.

[13] T. M. Cover and J. A. Thomas, *Elements of information theory*, Wiley-Interscience; 1991.

[14] B. J. Coi and L. Hanzo, ”Optimum Mode-Switching Assisted Adaptive Modulation,” *Proc. IEEE Globecom’01*, Nov. 6, 2001, pp. 3316-3320.

[15] J. M. Cioffi, *Lecture Notes for Advanced Digital Communications*, Standford, Fall 1997.

[16] A. Duel-Hallen, S. Hu and H. Hallen, "Long Range Prediction of Fading Signals: Enabling Adaptive Transmission for Mobile Radio Channels," *IEEE Signal Processing Mag.*, Vol. 17, No. 3, pp. 62 – 75, May 2000.

[17] T. Ericson, “ A Gaussian Channel with Slow Fading,” *IEEE Trans. Inform. Theory*, vol. IT-16, pp. 353-355, May 1970

[18] T. Eyceoz, A. Duel-Hallen, H. Hallen, "Deterministic Channel Modeling and Long Range Prediction of Fast Fading Mobile Radio Channels," *IEEE Commun. Lett.*, Vol. 2, No. 9, Sept. 1998, pp. 254 – 256.

[19] T. Eyceoz, S. Hu, and A. Duel-Hallen, "Performance Analysis of Long Range Prediction for Fast Fading Channels," *Proc. of 33rd Annual Conf. on Inform. Sciences and Systems CISS'99*, March 1999, Vol. II, pp. 656 - 661.

[20] T. Ekman and G. Kubin, “Nonlinear Prediction of Mobile Radio Channels: Measurements and MARS Modeling Designs,” *Proc. IEEE International Conference on Acustics, Speech and Signal Processing*, March 1999, pp. 2667-2670.

[21] A. J. Goldsmith and S.G. Chua, “Variable-Rate Variable-power MQAM for Fading Channels”, *IEEE Trans. Comm*, vol. 45, No 10, Oct 1997, pp1218-1230.

[22] A. J. Goldsmith and S. G. Chua, "Adaptive Coded Modulation for Fading Channels," *IEEE Trans. Commun.*, Vol. 46, No. 5, May 1998, pp. 595 - 601.

[23] A. J. Goldsmith and P. Varaiya, “Capacity of fading channel with channel side information,” *IEEE Transactions Information Theory*, vol: 43, Nov. 1997 pp. 1986 – 1992.

[24] D. L. Goeckel, “Adaptive Coding for Time-Varying Channels Using Outdated Channel Estimates,” *IEEE Trans. Commun*., Vol. 47, No. 6, June 1999, pp. 845-855.

[25] J. Hayes, “Adaptive Feedback Communications, “ IEEE Communication Technology, vol. 16, pp. 29-34, Feb. 1968.

[26] D. Hughes-Hartogs, Ensemble Modem Structure for Imperfect Transmission Media.

[27] J. K. Hwang, J. H. Winters, “Sinusoidal Modeling and Prediction of Fast Fading Processes,” *Proc. IEEE Globecom’98*, Nov. 1998, pp. 667-672.

[28] S. Haykin, *Adaptive Filter Theory, 3rd edition*, Prentice-Hall, 1996.

[29] S. Hu, H. Hallen, A. Duel-Hallen, "Physical channel modeling, adaptive prediction and transmitter diversity for flat fading mobile channel," IEEE SPAWC'99, pp. 387-390, May 1999.

[30] S. Hu, T. Eyeceos, A. Duel-Hallen, and H. Hallen, "Transmitter Antenna Diversity and Adaptive Signaling Using Long Range Prediction for Fast Fading DS/CDMA Mobile Radio Channels," *IEEE WCNC'99* Sep. 1999, pp. 824-828.

[31] S. Hu, A. Duel-Hallen, H. Hallen, "Long Range Prediction Makes Adaptive Modulation Feasible for Realistic Mobile Radio Channels," *Proc. of 34rd Annual Conf. on Inform. Sciences and Systems CISS'2000,* Vol. I, pp. WP4-7 ~ WP4-13, March 2000.

[32] S. Hu, A. Duel-Hallen, "Combined Adaptive Modulation and Transmitter Diversity Using Long Range Prediction for Flat Fading Mobile Radio Channels" *IEEE GLOBECOM'01*, vol. 2, Nov. 2001, pp. 1256-1261.

[33] H. Hallen, S. Hu, M. Lei and A. Duel-Hallen, "A Physical Model for Wireless Channels to Understand and Test Long Range Prediction of Flat Fading," *Proc. of WIRELESS 2001*, Calgary, Canada, July 9-11, 2001.

[34] H. Hallen, A. Duel-Hallen, S. Hu, T. S. Yang, M. Lei, "A Physical Model for Wireless Channels to Provide Insights for Long Range Prediction," *Proc. MILCOM'02*, vol. 1, pp.627-631, Oct. 7-10, 2002.

[35] W. C. Jakes, *Microwave Mobile Communications*, Wiley, New York, 1974.

[36] B. S. Krongold, "Computationally Efficient Optimal Power Allocation Algorithm for Multicarrier Communication Systems

[37] C. Kose, and D. L. Goeckel, "On Power Adaptation in Adaptive Signaling Systems," *IEEE Trans. Commun*., vol. 48, Nov. 2000, pp. 1769-1773.

[38] J. Lee, R. Arnott, K. Hamabe, and N. Takano, "Adaptive Modulation Switching Level Control in High Speed Downlink Packet Access Transmission," *3rd International Conference on 3G Mobile Communication Technologies*, May 2002, pp. 156-159.

[39] Y. Li, L. J. Cimini, and N. R. Sollenberger, "Robust Channel Estimation for OFDM Systems with Rapid Dispersive Fading Channel," *IEEE Trans. Commun*., vol. 46, pp. 902-915, Apr. 1998.

[40] S. Nanda, K. Balachandran, and S. Kumar, "Adaptive technique in wireless packet data serveice," *IEEE commun. Mag*., vol. 38, pp. 54-64, Jan. 2000.

[41] T. Ottosson, A. Svensson, “On Schemes for Multirate support in DS-CDMA Systems,” *Wireless Personal Communications*, vol. 6, pp. 265-287, March 1998.

[42] P. Ormeci, X. Liu, D.L. Goeckel, and R. D. Wesel, “Adaptive Bit-interleaved coded Modulation,” *IEEE Trans. Commun*. Vol. 49, Sep. 2001, pp. 1572-1582.

[43] A. Papoulis, *Probability, Random Variables, and Stochastic Process*, 3rd edition, McGraw-Hill, 1991.

[44] J. G. Proakis, “Advances in Equalization for Intersymbol interference,” *Advances in Communcation Systems*, vol. 4 A. J Viterbi(ed.), 1975, Academic, New York.

[45] J. G. Proakis, *Digital Communications*. Third Edition, McGraw-Hill, 1995.

[46] J. G. Proakis, and D. G. Manolakis, *Digital Signal Processing: Principle, Algorithms, and Application*. Third Edition, Prentice Hall, 1996

[47] T. Pollet, M. Van Bladel, M. Moeneclaey, ”BER sensitivity of OFDM Systems to Carrier Frequency Offset and Wiener Phase Noise,” *IEEE Trans. Commun.*, Vol. 43, No. 234, 2/3/4 1995, pp. 191-193.

[48] M. F. Pop, and N.C. Beaulieu, “Limitations of Sum-of-Sinusoids Fading Channel Simulators,” *IEEE Trans. Commun.*, vol. 49, No. 4, Apr. 2001, pp. 699-708.

[49] X. Qiu, and K. Chawla, ”On the Performance of Adaptive Modulation in Cellular Systems,” *IEEE Trans. Commun*. Vol. 47, Hune 1999, pp. 884-895.

[50] T. S. Rappaport, S. Y. Seidel, and K. Takamizawa, “Statistics Channel Model for Factory and Open Plan Building Radio Communication System design,” *IEEE Trans. Commun*., Vol. 39, No. 5, May 1991.

[51] T. S. Rappaport, *Wireless Communications: Principles and Practice*. Prentice-Hall, 1996.

[52] B. R. Salzberg, “Performance of an efficient parallel data transmission system,” *IEEE Trans. Comm*., Vol. 15, pp. 805-813, Dec. 1967.

[53] J. Salz, “Optimum Mean-Square Decision Feedback Equalization,” *Bell Syst. Tech. J*., Vol. 52 pp. 1341-1373,October.

[54] G. L. Stuber, *Principles of Mobile Communcarion*, Kluwer Academic Publishers, 2000.

[55] S. Sampei and T. Sunaga, “Rayleigh fading compensation for QAM in land mobile radio communications,” *IEEE Trans. Vehicular Technology*, vol. 42, pp. 137-147, May 1993.

[56] H. Su, and E. Geraniotis, “Adaptive Closed-loop Power Control with Quantized Feedback and Loop Filtering,” *IEEE Wireless Commun.* Vol 1, Jan. 2002, pp.76-86.

[57] T. Ue, S. Sampei, N. Morinaga and K. Hamaguchi, "Symbol Rate and Modulation Level-Controlled Adaptive Modulation/TDMA/TDD System for High-Bit-Rate Wireless Data Transmission," *IEEE Trans. Veh. Technol*, Vol. 47, No. 4, Nov. 1998, pp. 1134 - 1147.

[58] M. K. Simon and M. S. Alouini, *Digital Communication over Fading Channels*, 1st edition, John Wiley & Sons; 2000.

[59] M. K. Simon and Mohamed-Slim Alouini, “A Simple Single Integral Representation of The Bivariate Rayleigh Distribution,” *IEEE Commu. Letters*, Vol: 2 May 1998, pp. 128 –130.

[60] K. Skowratanont and J. A. Chambers," A Comparison of Blind Channel Estimation and Equalisation Techniques for A Fading Environment," *6th IEE Conference on Telecommunications*, No. 451, March 1998, pp. 27-31.

[61] L. Thibault, M. T. Le, “Performance Evaluation of COFDM for Digital Audio Broadcasting. I.: Parametric Study,” *IEEE Trans. Broadcasting*, Vol. 43, No. 1, pp. 64-75, March 1997.

[62] M. Torabi, and M.R. Soleymani, “Variable-rate OFDM Systems with Selective Antenna Diversity and Adaptive Modulation,” IEEE Vehicular Tech. Conference, vol. 1, Apr. 2003, pp. 562-566.

[63] S. B. Weinstein and P.M. EBERT, “Data Transmission by Frequency-Division Multiplexing Using the Discrete Fourier Transform,” *IEEE. Trans. Commun,* vol. 19, No 5, Oct 1971, pp628-634.

[64] W. Webb and R. Steele, “Variable rate QAM for Mobile Radio,” *IEEE Trans. Comm.*, vol. 43, pp. 2223-2230, July 1995.

[65] H. Zhu, and K.J.R. Liu, “Joint Adaptive Power and Modulation Management in Wireless Networks with Antenna diversity,” *Proc. Sensor Array and Multichannel Signal Processing Workshop*, Aug. 2002, pp. 278-282.

[66] S. Zhou and G. B. *Giannakis*, "Adaptive Modulation for Multi-Antenna Transmissions with Channel Mean Feedback," *IEEE ICC'03*, Vol. 4, 2003, pp. 2281 -2285.